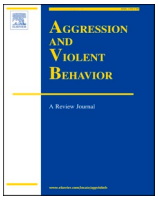
Aggression and Violent Behavior 68 (2023) 101798

Contents lists available at ScienceDirect

Aggression and Violent Behavior

journal homepage: www.elsevier.com/locate/aggviobeh



Parent training programs for preventing and treating antisocial behavior in children and adolescents: A comprehensive meta-analysis of international studies

Andreas Beelmann *, Louisa S. Arnold, Judith Hercher

Institute of Psychology, Friedrich-Schiller-University of Jena, 07743, Germany

ARTICLE INFO

Keywords:
Parent training
Meta-analysis
Antisocial behavior
Children
Adolescents

ABSTRACT

This article presents a comprehensive meta-analysis of international studies on the effects of parent training programs (PTP) on antisocial behavior (ASB) in children and adolescents. From systematic literature searches of 7219 reports, we finally selected 239 eligible reports with 241 independent studies and 279 comparisons between a program and a control condition up to the publication year 2020. Although most interventions were based on a cognitive-behavioral approach, we also found a great variety of programs and applications. Overall, the mean effect for PTP was positive for parent/family and ASB outcomes ($d = 0.46$ and $d = 0.47$, respectively using the random effect model at postintervention). We also found higher effects on more proximal parental outcomes such as parental stress, parental competencies, and parent–child interaction/relation. However, more distal outcomes such as marital satisfaction or parent psychopathology revealed lower effect sizes. In addition, the link between changes in parental/family outcomes and changes in ASB was significant across several outcome types, thus confirming the general causal assumption of PTP. Postintervention effects were stable across several moderators, although clinical applications revealed slightly higher effect sizes than preventive applications. Several findings cast some doubt on these generally positive results: For example, effect sizes decreased considerably in not only short- (3 to 12 months) but also especially long-term follow ups (12 months or more), and the vast majority of outcome assessments stemmed from parent ratings. Finally, we found a clear negative connection between sample and effect size. Whether this is due to publication bias or indicates a better implementation quality in smaller studies remains an open question.

Antisocial behavior (ASB) in children and adolescents such as aggression, violence, delinquency, and crime is a serious problem that probably affects every society in the world (Belfer, 2008). Moreover, ASB is one of the most prevalent behavior problems in children and adolescents with 10 to 20 %—depending on the assessment method, the informant, and the selected sample—exhibiting serious problems at least temporarily in their social development (e.g., Frick, 1998; Lahey et al., 1999), and clinical forms of these behavior problems such as oppositional defiant disorder and conduct disorder (see American Psychiatric Association, 2013) making up the vast majority of childhood referrals to psychological and psychiatric services (Maughan et al., 2004; National Institute for Health and Care Excellence, 2013). Of course, ASB extends along a continuum of severity ranging from disruptive behavior that burdens parents, teachers, and peers up to highly problematic violent and criminal behavior that can cause serious harm to others (Burke et al., 2002; Frick et al., 1993; Loeber et al., 2000). In addition, the long-term prognosis for these behavior problems, especially if left untreated, can include lower levels of educational attainment, unemployment, drug and alcohol abuse, crime, and incarceration along with higher levels of school dropout, hospitalization and mortality, family breakdown, and, finally, the intergenerational transmission of ASB (Farrington, 2007; Lipsey & Derzon, 1998; Moffitt, 1993; Moore et al., 2017). These long-term problems related to ASB in childhood and youth result in extensive costs to health, social, education, and legal services. It has been estimated that the costs for individuals with a clinically diagnosed conduct disorder, for example, are 10 times higher than for children without such a diagnosis. Costs for nonclinical ASB have been found to be 3.5 times higher than for children and youth not displaying such behavior problems (Fergusson et al., 2005; Scott et al., 2001). In addition, the cost of criminal behavior for societies is remarkable. Systematic

* Corresponding author at: Humboldtstr, 26, 07743 Jena, Germany.
E-mail address: andreas.beelmann@uni-jena.de (A. Beelmann).

https://doi.org/10.1016/j.avb.2022.101798
Received 7 December 2020; Received in revised form 15 September 2022; Accepted 23 October 2022
Available online 27 October 2022

analyses have shown that individuals with ASB produce a cumulative societal cost in the millions (Muñoz et al., 2004; Romeo et al., 2006). Hence, ASB is a serious and very prevalent behavior problem that needs to be addressed and, at best, prevented because of its severe consequences and high societal cost.

Parenting and child-rearing behavior as possible causes of ASB have been discussed intensively during the last decades (e.g., Farrington et al., 2012). The ways in which parenting affects child development and behavior have been captured in a number of developmental models. Most central to these are attachment theory (Bowlby, 1969, 1973) and research on parenting styles (Baumrind, 1966; Maccoby & Martin, 1983). Both place parents in the pivotal position of being the most important agents of socialization, especially for young children. More generally, how parents shape their children's development and behavior is spelled out in social learning theory (Bandura, 1977): Children learn from their parents by relating to them, imitating them, and reacting to their behavior and attitudes. Ineffective parenting skills such as low levels of parental supervision and involvement as well as punitive and inconsistent discipline result in dysfunctional learning situations. This means that adverse child behavior is positively reinforced, whereas positive behavior becomes less frequent because it is not attended to appropriately (Reid et al., 2002). At the same time, nonetheless, it is also important to recognize that the relationship between parenting and child behavior is bidirectional: For example, temperamental differences in children's behavior might also shape the ways parents habitually react to their children. Thus, negative child behavior and ineffective parenting skills can combine into coercive cycles in which mutual hostility and aggression develop and consolidate over time (Granic & Patterson, 2006; Patterson et al., 1992).

A great deal of developmental and clinical research has shown that ASB has multiple causes and can be explained only through complex developmental models integrating biological, psychological, and social risk or protective factors (e.g., Beelmann, 2012; Farrington, Gaffney, Ttofi, et al., 2017; Granic & Patterson, 2006). These models acknowledge the multicausality of ASB as well as the antagonistic and accumulative relations between the variables related to its development. Nonetheless, numerous empirical studies confirm that deficits in parenting behavior such as poor monitoring, parental hostility, and dysfunctional discipline are among the most influential risk factors for ASB in child and youth development (Braga et al., 2017; Forehand et al., 2013; Hoeve et al., 2009; Rothbaum & Weisz, 1994). Indeed, a meta-analysis of >1400 publications confirmed that several parenting characteristics or parenting styles are associated consistently with ASB in children and adolescents (Pinquart, 2017). For example, parental warmth, behavioral control, and authoritative parenting style showed a consistently negative correlation to ASB (r ranging from −0.16 to −0.19), whereas harsh control, permissive, neglectful, and an authoritarian parenting style showed positive associations (r ranging from 0.19 to 0.22). These covariations could also be confirmed after controlling for bidirectional effects via cross-lagged and longitudinal analyses (see also reviews on longitudinal research by Farrington, Gaffney, Ttofi, et al., 2017; Jolliffe et al., 2017). Consequently, it is quite clear that efforts to prevent and treat ASB should focus on the parents or other primary caretakers in order to foster their parenting skills as the crucial leverage points of change.

1. Parent training programs as prevention and intervention for ASB

Over the past 50 years, various parent training programs (PTP) have been developed and widely applied as measures for preventing and treating ASB in children and adolescents. Today, there are numerous programs and several theoretical models that reveal different histories within the behavioral (Patterson, 1982), humanistic (Gordon, 1970), and psychodynamic (Adlerian) traditions (Dreikurs & Soltz, 1964). Despite these different roots, PTP can generally be defined as curriculum- or manual-based interventions designed to promote parenting skills such as parenting supervision, nonaggressive limit setting, positive reinforcement, emotional and social support, perceiving and adequately interpreting a child's verbal and nonverbal communication, promoting the parent–child relationship, and delivering age-appropriate promotion of a child's development (see Briesmeister & Schaeffer, 2007; Kazdin, 2005, van Ryzin et al., 2016, for reviews). Several brand-name programs based on different theoretical approaches and varying in content and breadth (e.g., Incredible Years, Parent-Child-Interaction Therapy, Parent Management Training, Parent Effectiveness Training, Systematic Training in Effective Parenting, Triple P) have been developed and published in recent decades, thus making this approach one of the most widely applied prevention and intervention measures within the area of ASB (see Beelmann & Raabe, 2009; Weber et al., 2019).

In general, PTP aim to prevent or reduce ASB in children and adolescents by improving parenting skills and behavior. Several alternative models depict how increased parenting skills prevent or effectively lower ASB (Forehand et al., 2014; Sandler et al., 2011). One general behavioral model assumes that PTP not only promote positive parenting behavior and parenting self-efficacy but also impact directly on children's behavioral development (including problem behavior such as ASB). For example, the theory of coercive interaction (Granic & Patterson, 2006; Patterson et al., 1992) assumes that inadequate limit setting in parent–child interactions during problem situations establishes a vicious circle and a system of negative and forced interactions between parents and children that leads finally to the development of ASB in early childhood. However, other models assume more complex causal pathways by plotting how increased parenting skills lead to a positive social development and the prevention or reduction of problem behavior (see Beelmann & Raabe, 2009; Forehand et al., 2014; Sandler et al., 2011, for reviews). For example, Sandler et al. (2011) have outlined a youth–environment transactional model that assumes that parenting changes lead to more positive transactions between young persons and their environment (e.g., more positive reactions, lower rejections by peers), and these finally result in lower problem behavior and greater developmental competencies. Empirically, longitudinal research supports the prominent role of effective discipline as a mediator between PTP and either preventing or decreasing ASB (Forehand et al., 2014) with a coercive parenting style implying the highest risk for ASB (LoBraico et al., 2020) and supportive parenting protecting from antisocial development (van Heel et al., 2019).

2. Prior reviews of parent training programs

In the past decades, PTP have been evaluated intensively and numerous reviews and meta-analyses have summarized research on PTP and related topics (see Beelmann & Raabe, 2009; Farrington, Gaffney, Ttofi, et al., 2017; Lösel & Bender, 2016; Matjasko et al., 2012; Mingebach et al., 2018, Sandler et al., 2011, Weber et al., 2019; for reviews of reviews). However, all reviews either have several limitations or differ in breath and scope from the present review. Some reviews are outdated and therefore of limited value (e.g., Cedar & Levant, 1990; Serketich & Dumas, 1996). Others are restricted to special brand-name programs such as Triple P, PCIT, or Incredible Years (e.g., de Graaf, Speetjens, Smit, de Wolff, Tavecchio, 2008a, 2008b; Menting et al., 2013; Nowak & Heinrichs, 2008; Sanders et al., 2014; Thomas et al., 2017; Wilson et al., 2012) or special target groups like foster parents (Schoemaker et al., 2020). Some reviews are restricted to special age groups (e.g., Barlow et al., 2005; Dretzke et al., 2009; Piquero et al., 2008, 2009; Piquero et al., 2016) or analyze only long-term effects (van Aar et al., 2017). Further reviews have summarized results of clinical applications of PTP alone (Dretzke et al., 2005, 2009; Furlong et al., 2012; Maughan et al., 2005; McCard et al., 2006), others have integrated research on PTP together with other family-related interventions (Farrington & Welsh, 2003; Piquero et al., 2016; Weiss et al., 2015), are restricted to specific

settings and countries (see, e.g., Knerr et al., 2013, for PTP in low- and middle-income countries; Weiss et al., 2015, for studies conducted in Germany), or examined the influence of only selected outcome moderators (e.g., Leijten et al., 2013; Leijten et al., 2019; Lundahl et al., 2006; Lundahl et al., 2008; Reyno & McGrath, 2006; Wyatt Kaminski et al., 2008). Finally, all cited international meta-analyses are restricted to English-language publications, despite a growing body of international non-English literature such as German-language studies.

Most importantly, none of the cited reviews performed a systematic analysis of different outcome variables for ASB and parent- and family-related outcomes—although this would be useful in differentiating the effects of PTP (Weber et al., 2019). Hence, a more detailed and comprehensive analysis of the effectiveness of PTP is needed to identify the most promising programs and the most important moderators of outcomes. Therefore, the present meta-analysis extends existing reviews by (a) delivering a recent and comprehensive review on the outcomes of parent training programs for preventing and treating antisocial behavior problems in children and adolescents; (b) collecting all international studies without being restricted to English-language publications; (c) considering the whole age range of applications in childhood and adolescence (up to age 18); (d) including prevention programs and clinical treatments and all types of parent training programs; (e) covering the whole range of outcomes (e.g., antisocial behavior, parent and family outcomes); (f) providing an in-depth focus on methodological moderators and the effects of conflict of interest; and (g) encompassing a broader range of study designs by including RCTs and also high-quality quasiexperimental designs.

3. Method

3.1. Criteria for including and excluding studies

3.1.1. Types of study design

The review considered published and unpublished studies on the evaluation of a PTP using RCTs and high-quality quasiexperimental designs with at least two assessment points (pre–post) *and* two experimental groups (PTP condition vs. a comparable control condition). Comparison conditions were untreated groups or treatment-as-usual (TAU) or service-as-usual (SAU) conditions, as long as these were not (a) a fully described alternative program, (b) an intensive treatment (or treatment combination) that was applied to >50 % of the comparison group and tested against the outcome of the PTP condition, or (c) PTP was additionally delivered to a full treatment. We determined the comparability of the comparison group in non-RCT as well RCT designs according to demographic factors and the pretest level of central ASB outcome measures. Studies were excluded when comparison groups differed significantly on more than one demographic factor (e.g., age, gender, social class) or on antisocial behavior measures at pretest (significant differences or more than one *SD* difference). If studies applied PTP within multitreatment arm studies, results on the effectiveness of the PTP alone could be included if reported separately and if this was the only treatment this subsample received. Minimum sample size was 10 for each experimental group (PTP and control group). Although there is no precise threshold for a sufficient sample size in meta-analytic applications, we expected to have statistical parameters with acceptable statistical power from a total minimum sample size of 20. In addition, by excluding smaller samples, the meta-analysis would be less likely to be affected by publication bias (Dickersin, 1990; Gilbody et al., 2000; Newcombe, 1987; Thornton & Lee, 2000).

3.1.2. Types of interventions

The review integrated all studies on standardized (i.e., curriculum-based or manualized) PTP with the main aim of promoting parenting skills in order to prevent or treat ASB in children and adolescents. There were no restrictions regarding content, setting, and other modalities. In addition, the PTP could have been conducted with different target persons (mothers, fathers, or both parents) who did not necessarily have to be the parents as long as they were currently in charge of raising and parenting a child (e.g., foster parents, grandparents, other relatives). We excluded interventions with open formats (e.g., parent counseling) and programs focusing primarily on either promoting parent or family relationships (e.g., relationship education, family therapy) or teaching parenting skills in terms of physical care and nutrition within the first years of life. We also excluded studies testing combinations of PTP with other interventions (e.g., child training, early intervention programs). However, we included studies in which children took part in the parent training as in Parent–Child Interaction Therapy (Eyberg & Boggs, 1998).

3.1.3. Types of participants

All types of samples were included as long as the primary aim of the intervention was to prevent and treat ASB in children and adolescents up to the mean age of 18. There was no restriction according to demographic factors, specific risk factors, or intervention settings. However, the target groups recruited for the intervention could be on different levels of risk, meaning that the preventive approach was either universal (no particular risk), selective (some risk factors related to the outcome were present—e.g., parents from a problematic neighborhood), or indicated (first signs of the problematic outcome were present—e.g., children with high rates of externalizing behavior). In addition, we included clinical applications of PTP as a psychotherapeutic intervention containing groups with already established diagnoses (e.g., oppositional defiant disorder, conduct disorder, juvenile offenders) but excluded PTP with a primary focus on the prevention and treatment of other childhood and adolescent problems such as drug use, depression, or ADHD as well as programs especially designed for the prevention and treatment of child maltreatment and abuse. Finally, we also excluded studies that were directed toward parents facing certain life events (divorced mothers). However, if the main aim of these studies was to prevent or treat ASB in children and adolescents, these programs were included (coded as at-risk groups).

3.1.4. Types of outcome measures

The review summarized only studies containing at least one measure of ASB in the broadest sense (i.e., data on aggression, violence, delinquency, crime, and other forms of antisocial behavior) at one or more assessments after the termination of the program. General behavior problems scores were taken only if the scale predominately contained ASB items (e.g., inclusion of the ECBI but exclusion of the CBCL total problem score). Therefore, the primary outcome category was all variables assessing ASB across different types of behavior (e.g., aggression, criminal behavior), different assessment methods (behavioral ratings or observations), and different informants (parent or teacher ratings). Secondary outcomes were all parent- and family-related outcomes (e.g., parenting competencies, parent–child interaction, etc.) For mediational analyses, we assumed that effects on ASB outcomes would be caused by effects on parent- and family-related outcomes.

3.2. Search strategy

We searched for all published and unpublished studies in English or another main European language (e.g., German, Spanish, Italian, Dutch) up to the year 2020. We used the following four search strategies: (a) We conducted electronic searches in bibliographic databases and especially in PsycINFO, PSYNDEX, PsychARTICLES, Criminal Justice Abstracts, Web of Science, ERIC, PubMed, ProQuest Dissertations and Theses, Cochrane Central Register of Controlled Trials (CENTRAL), and additionally Google and Google Scholar. The keywords used for the electronic databases covered three generic search categories: parent training, antisocial behavior, and effectiveness. The proposed keywords for these categories are listed in Supplement A. (b) We compiled a list of brand names of established PTP (e.g., Incredible Years, Parent Effectiveness Training, Parent Management Training, Triple P) and searched

in further electronic databases. (c) We screened references in existing reviews or meta-analyses (see overview above) for primary studies. (4) We conducted a backward as well as a forward search with the final pool of included publications. In the backward search, we screened the reference lists of all included studies for further relevant publications. In the forward search, we screened all studies citing the included studies.

Using these strategies, we identified a total of 7219 research reports, but excluded a large percentage on the basis of the title and the abstract (5678). However, 1541 reports were checked with full-text analyses finally leading to 239 eligible research reports including 241 separate independent studies and a total of 279 comparisons between a PTP condition and a control condition (see flowchart in Fig. 1 and references in Supplement B). The main reasons for exclusions were: (a) no or no adequate/comparable control groups (27.2 % of all excluded studies with full-text analysis), (b) missing outcome on ASB (16.0 %), (c) no evaluation study (14.2 %), (d) a double/supplemental publication (11.1 %), and (e) no PTP as intervention (10.0 %) (see Fig. 1).

3.3. Coding procedures

The three authors and a total of five trained assistants coded all reports, studies, and comparisons according to the characteristics of the report (e.g., year and country of origin), methods (e.g., design and follow-up assessments), intervention programs (e.g., type, intensity, or setting), and the children and adolescents trained (e.g., age, gender, risk factors, and level of intervention). On the effect-size level, we calculated the effect sizes and coded the outcome categories of ASB (aggression, delinquency, oppositional-defiant behavior, and general measures of antisocial behavior) and parent/family outcomes (PFO; e.g., parental competencies, parental cognitions, parental stress, parent–child interaction, or family conflicts), the assessment method (self-report, parent and teacher ratings, and official data), and the time between the termination of the intervention and the outcome assessment. All posttests and follow ups were included regardless of duration. However, to unify quantitative analyses, we applied a standardized time metric: All measurements up to 3 months after the termination of training were summarized as postintervention effects; all measurements between 3 and up to 12 months, as short-term follow-up effects; and all measurements after 12 months and more, as long-term follow-up effects. To ensure reliability, the coding team met at regular intervals to discuss unclear issues until all intercoder reliabilities attained Cohen's $\kappa > 0.8$ (attainted by the independent rating of 20 comparisons).

3.4. Statistical analyses

We calculated unbiased standardized mean differences from means and standard deviations of treatment and control groups or from test statistics as recommended by Lipsey and Wilson (2001). We used pooled pretest standard deviations for effect-size calculation as proposed by Morris (2008). If the reports mentioned nonsignificant results with no further details, we counted these as zero effects. Because many studies contained several relevant outcome measures, we computed 965

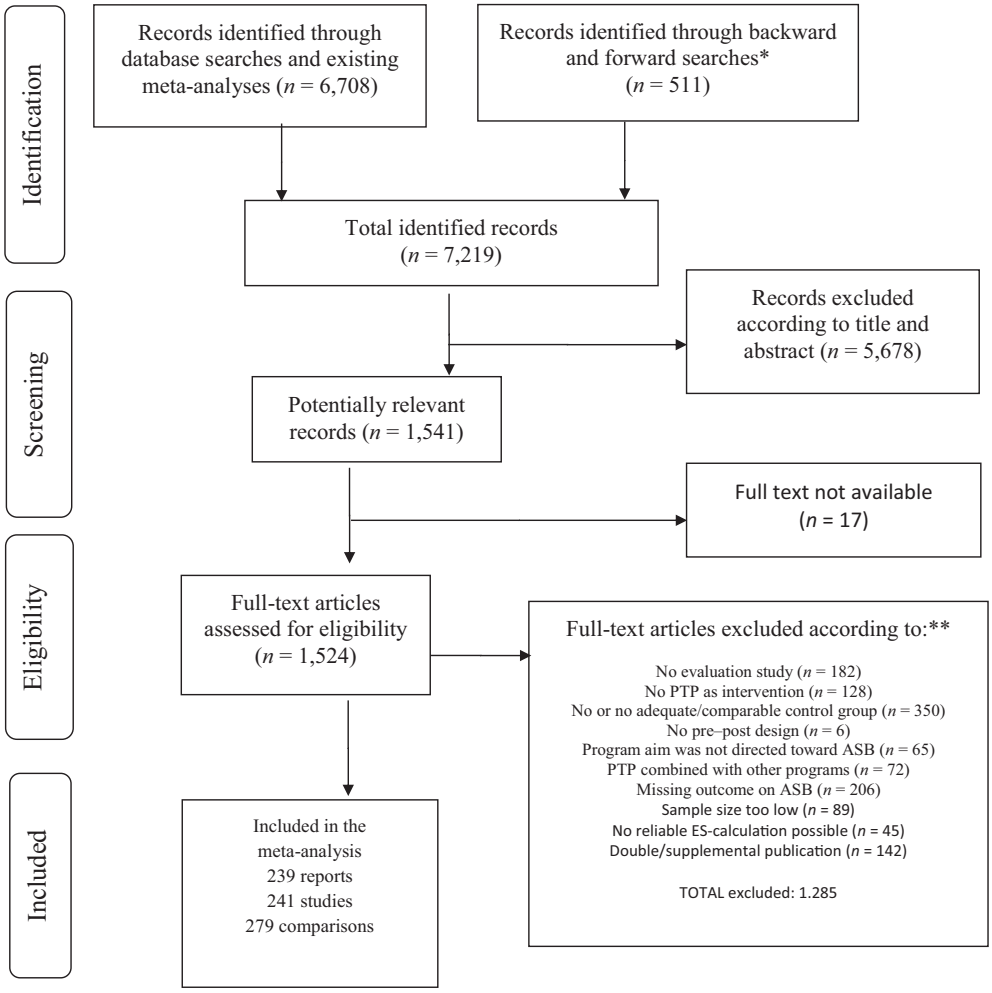


Fig. 1. Prisma flow chart for identifying relevant research reports, studies, and comparisons.
* Searches only within included reports. ** In the order of exclusion.

individual effect sizes for the various ASB outcomes (702 for postintervention, 147 for short-term follow-up intervals, and 116 for long-term follow-up intervals) and 1689 individual effect sizes for PFO (1253 for postintervention, 277 for short-term follow-up intervals, and 159 for long-term follow-up intervals). Most of these effect sizes were calculated from means and standard deviations (85.9 %), 11.5 % were recalculated from reported test statistics, 1.8 % were taken directly from the reports, and 0.9 % were estimated as zero effects. All effect sizes were polarized in the direction of program effects—for example, a reduction in antisocial behavior was displayed as a positive effect size.

We then aggregated these individual effect sizes on the comparison level separately for ASB and PFO by averaging them within each comparison according to the coded postintervention and follow-up time intervals (see above). Thus, we calculated only one effect size for ASB and PFO for each time interval and each comparison. In total, within the 279 comparisons, we could conduct 349 outcome assessments for ASB (258 at postintervention, 57 at short-term follow up, and 34 at long-term follow up) and 304 outcome assessments for PFO (228 at postintervention, 51 at short-term follow up, and 27 at long-term follow up; see Table 1). We than integrated comparison effect sizes across comparisons following Lipsey and Wilson's (2001) recommendations on weighting by the inverse of the squared standard error using the random effect size model developed by Hedges and Olkin (1985). For the calculation of effect sizes and their subsequent integration, we used SPSS 25 with macros developed by Wilson (2010); for the design of the funnel plot and the calculation of the trim-and-fill analysis, the procedure Metafor for R.

Table 1
Description of parent training programs.

	Number of comparisons	Percent
Intensity		
Up to 5 h	37	13.3
5 to 10 h	53	19.0
10 to 15 h	67	24.0
16 to 20 h	60	21.5
>20 h	58	20.8
Not specified	4	1.4
Format		
Group	135	48.4
Individual	58	20.8
Group and individual	50	17.9
Self-help (books, videos)	30	10.8
Not specified	6	2.2
Setting		
At home	41	14.7
Clinic	30	10.8
Community facility	31	11.1
Kindergarten/School	28	10.0
University	10	3.6
Mixed	38	13.6
Not specified	101	36.2
Administrator		
Study author	35	12.5
University staff (not author)	81	29.0
Professional practitioner	65	23.3
Mixed	27	9.7
None	27	9.7
Not specified	44	15.8
Target group		
Parents[a]	177	63.4
Mothers[b]	78	28.0
Fathers[c]	4	1.4
Parent/Mother–child	11	3.9
Foster/Step parents	9	3.2

[a] Included all unspecified cases and cases with mothers and fathers (at least >10 %) as participants.
[b] Included all cases with mothers only (>90 %).
[c] Included all cases with fathers only.

4. Results

4.1. Description of the dataset

Most *reports* had been published as journal articles (89.5 %), followed by dissertations (8.4 %), unpublished research reports (1.3 %), and books or book chapters (0.8 %). Publication date varied between 1974 and 2019, with most reports appearing between 2011 and 2020 (94 or 39.3 % of reports). A total of 9.6 % (23) were published up to 1990; 16.3 % (39), from 1991 up to the year 2000; and 34.7 % (83), from 2001 to 2010.

Most *comparisons* were conducted in the United States ($k = 119$, 42.7 %) and Australia ($k = 50$, 17.9 %). Further countries with at least 10 comparisons were the United Kingdom ($k = 20$, 7.1 %), Sweden ($k = 13$, 4.7 %), Germany ($k = 12$, 4.3 %), Canada, and the Netherlands (each $k = 10$, 3.6 %). Further comparisons were conducted in Ireland ($k = 7$, 2.5 %); Norway ($k = 6$, 2.2 %); Switzerland, Belgium, and China (each $k = 4$, 1.4 %); Spain ($k = 3$, 1.1 %); New Zealand, Iran, Indonesia, Portugal, and Finland (each $k = 2$, 0.7 %); and, finally, Japan, Burundi, Israel, Iceland, Poland, Pakistan, and Rumania (each $k = 1$, 0.4 %).

Table 1 presents the characteristics of the evaluated parent training programs according to intensity, format, setting and administrator, and target group. In general, data on the intensity (total hr of program) showed a considerable variation in PTP (from 1 to 36 h), and that it was delivered predominantly in a group format within different settings conducted by several administrators. The target groups were mainly both parents, although a lot of studies do not report the participation of mothers and fathers separately. If they do, then it was mostly mothers who participated. Overall, a lot of information about these characteristics was not specified in the reports.

4.2. Mean effect size on ASB and PFO at postintervention and follow ups

Table 2 summarizes the overall results of PTP on ASB and PFO for the different assessment intervals.

Overall, we found nearly the same substantial effect sizes for both ASB and PFO at postintervention ($d_+ = 0.47$, and $d_+ = 0.46$, respectively), but also decreasing effect sizes for both outcomes at short-term and long-term follow ups ($d_+ = 0.22$ and 0.12 for ASB and $d_+ = 0.27$ and 0.13 for PFO). However, all overall effect sizes differed significantly from zero, indicating systematic improvements across studies, outcome assessments, and assessment times. In addition, effect sizes for both outcome areas were heterogeneous at postintervention and short-term follow up, but homogeneous at long-term follow up.

4.3. Analysis of publication bias

Analyses of publication bias were conducted separately for PFO and ASB outcomes on the basis of postintervention results. For PFO outcomes, a linear regression from sample size to the unweighted effect sizes indicated a significant publication bias ($\beta = -0.29$, $t = -4.49$, $p < .001$ see Macaskill et al., 2001) and a significant Begg's rank correlation ($r_{tau} = -0.26$, $z = -5.76$, $p < .000$, see Begg & Mazumdar, 1994). The funnel plot (see Fig. 2) with a subsequent trim-and-fill analysis (Rothstein, 2008) yielded an estimated 26 studies on the left side with a corrected mean effect size of $d_+ = 0.38$ ($z = 11.92$, $p < .001$). The test for funnel plot asymmetry also delivered a significant result ($t = 7.99$, $df = 224$, $p > .001$; see Egger et al., 1997).

ASB outcomes also revealed a systematic publication bias, with a significant regression from sample sizes to the unweighted posttest effect sizes ($\beta = -0.34$, $t = -5.78$, $p < .001$. Begg's rank correlation was also significant ($r_{tau} = -0.31$, $z = -7.47$, $p < .001$). A funnel plot (see Fig. 3) with a subsequent trim-and-fill analysis yielded the substantial estimation of 21 missing studies on the left side with a corrected mean effect size of $d_+ = 0.42$ ($z = 15.05$, $p < .001$). This finding was further supported by a significant regression test for funnel plot asymmetry ($t =$

Table 2
Overall effect size statistics for parent/family outcomes (PFO) and antisocial behavior outcomes (ASB).

Assessment period	d_+	±CI	PFO *k*	%I	*OR*	d_+	±CI	ASB *k*	%I	*OR*
Postintervention	0.46***	±0.05	226	11.2	2.32	0.47***	±0.05	258	11.4	2.34
	Q_t (225) 601.68***, $I^2 = 62.6$					Q_t (257) = 735.34***, $I^2 = 65.1$				
Short-term FU	0.27***	±0.06	51	6.5	1.62	0.22***	±0.07	57	5.5	1.50
	Q_t (50) = 79.51**, $I^2 = 37.1$					Q_t (56) = 132,32***, $I^2 = 57.7$				
Long-term FU	0.13***	±0.06	27	3.2	1.27	0.12**	±0.07	34	3.0	1.24
	Q_t (26) = 23.47, $I^2 < 0.$					Q_t (33) = 63.07, $I^2 = 47.7$				

Notes. FU = follow up. All effect sizes were calculated according to the random effect size model. d_+ = weighted mean effect size; CI = limits of the 95 % confidence interval; *k* = number of comparisons; %I = improvement in percent; *OR* = odds ratio; I^2 = percentage of variance due to heterogeneity; *Qt* = homogeneity statistic.
** $p < .01$.
*** $p < .001$.

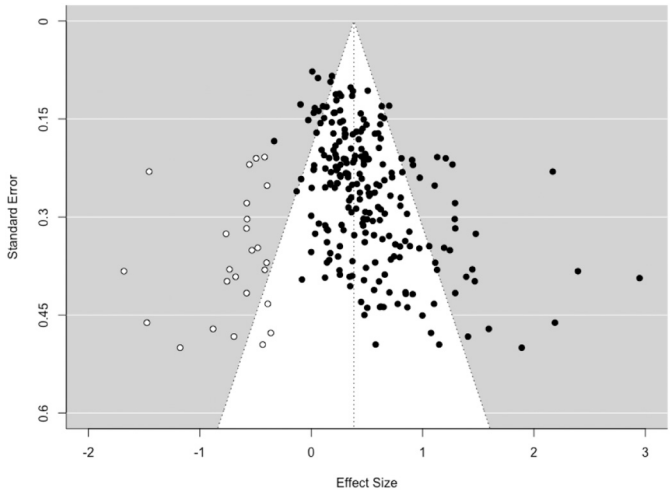


Fig. 2. Funnel plot and trim-and-fill analysis for parent and family outcomes.

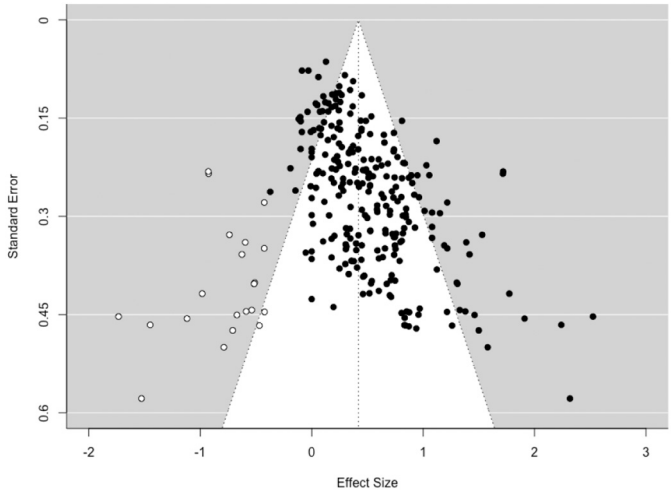


Fig. 3. Funnel plot and trim-and-fill analysis for antisocial behavior outcomes.

11.33, $df = 256$, $p < .001$).

4.4. Effects on different types of outcomes

Table 3 shows the effect sizes for different categories of PFO and ASB split by the different assessment periods. Overall, we found a general tendency for decreasing effects with increasing follow-up period across all outcome variables. Effect sizes on ASB were generally significant for the postintervention period with slightly lower values for aggressive and delinquent behavior. Significant follow-up effects could be confirmed only for oppositional/disruptive behavior and general measures of ASB (e.g., general ASB scales such as ECBI) that was by far the most frequent category.

Table 3
Effect sizes of different categories of PFO and ASB.

PFO/ASB categories	POST		ST-FU		LT-FU	
	d_+	*k*	d_+	*k*	d_+	*k*
Parental/family outcomes						
Parental competence	0.46*+	154	0.31*+	42	0.19*	20
Parental cognitions	0.52*+	91	0.28*	13	0.22	7
Parental stress	0.43*+	104	0.19*+	17	0.01	7
Parent–child interaction/ relation	0.65*+	69	0.27*+	16	0.11*	14
Aversive parental behavior	0.43*+	68	0.19*+	13	0.17*+	9
Parent relation/Family quality	0.24*+	46	0.24*	8	0.05	6
Parental psychopathology	0.25*+	57	0.18*+	17	0.05	4
Antisocial behavior outcomes						
Oppositional/Disruptive behavior	0.41*+	32	0.16*+	13	0.14*	8
Aggressive behavior	0.30*+	35	0.10+	7	0.07+	5
Delinquent behavior	0.28*+	15	−0.05	6	0.06	7
General ASB	0.47*+	238	0.26*+	47	0.13*+	32

Notes. POST = postintervention; ST-FU = short-term follow up; LT-FU = long-term follow up; *d*+ = weighted mean effect size; *k* = number of comparisons; * effect size differed significantly from zero; + = effect size showed significant heterogeneity.

Effect sizes were relatively equal across the different categories of PFO. However, there were some exceptions to this general pattern: For example, effects on parent–child interaction/relation were higher than for the other PFO categories and had at least some significant follow-up effects. In contrast, results on parent relation/family quality (e.g., marital satisfaction, family cohesion) as well as on parental psychopathology (e.g., maternal depression) were somewhat lower compared to more proximal (primary) outcomes on parental competence, cognitions, stress, and aversive behavior. Measures of parental competence and of aversive parental behavior showed small but significant follow-up effects even after one year or more (long-term follow-up period).

It was remarkable to see that outcomes also varied by assessment method and informants. In general, we found various assessment tools, but only a very restricted outcome assessment. For example, the vast majority of assessment methods were rating scales (80.8 % of all individual effect sizes) with the Eyberg Child Behavior Inventory (ECBI) and the externalizing subscale of the Child Behavior Check List (CBCL) as the most frequently applied instruments. A further 13.5 % of the individual effect sizes came from behavioral observations and 4.8 % from interviews. Only four individual effect sizes (0.2 %) were calculated from official data (school and police records) and a further 0.7 % were composite scores based on mixed methods. The same uneven distribution was obtained when looking at the informant. A total of 52.0 % of all

effect sizes were based on parent's (gender unspecified), a further 18.6 % on mother's, and 6.6 % on father's information resulting in nearly four out of five effect sizes representing estimates by the target group. Further informants were independent observers/interviewers (14.3 %), teachers (3.7 %), and the child/adolescent her or himself (4.4 %).

On the basis of these distributions, it was not surprising that the mean effect size at postintervention for the most common assessment methods and informants was around the mean of ASB and PFO outcomes ($d_+ = 0.43$, $k = 249$ for rating scales as assessment method, and $d_+ = 0.45$, $k = 243$ for parents as informants at postintervention). However, observations ($d_+ = 0.64$, $k = 63$) as well as independent observers and interviewers revealed higher effect estimations ($d_+ = 0.72$, $k = 60$). Most remarkable was the lower effect sizes for teacher as informant ($d_+ = 0.08$, $k = 37$, ns) as well as for self-reports by the child/adolescent ($d_+ = 0.14$, $k = 24$, $p < .05$), both based totally on an ASB assessment.

Finally, we tested the mediational assumption that effects on PFO would predict effects on ASB. Weighted meta-regressions yielded a substantial prediction from the strength of overall PFO to the level of ASB outcomes ($\beta = 0.55$, $k = 217$, $p < .001$). The highest connections were found for parental stress ($\beta = 0.63$, $k = 104$, $p < .001$) and parent–child interaction/relation ($\beta = 0.57$, $k = 69$, $p < .001$). There were medium predictions for parental competence ($\beta = 0.44$, $k = 153$, $p < .001$) and aversive parental behavior ($\beta = 0.47$, $k = 68$, $p < .001$), and lower and partially nonsignificant predictions for parental cognitions ($\beta = 0.32$, $k = 91$, $p < .01$), relationship/family quality ($\beta = 0.23$, $k = 46$, *ns*), and parent psychopathology ($\beta = 0.17$, $k = 57$, *ns*).

4.5. Effect sizes according to program type, level of intervention, and age

Table 4 reports the effect sizes according to several effect moderators. First, we compared different program types with their theoretical approaches and those brand-name programs that had at least ten comparisons. In general, all programs had at least medium effects on PFO as well as on ASB with widely comparable effect sizes, although a between test of heterogeneity led to significant differences between categories on PFO and ASB. However, these results were due to a few exceptions (see Table 4). Effect sizes on PFO and ASB were significantly higher for Parent-Child-Interaction Therapy (PCIT), humanistic programs, and, to a lower extent, Adlerian programs. However, these programs were generally based on a smaller data base then other program types. Otherwise, we found that most program categories exhibited homogeneity, at least for ASB outcomes, indicating a low variability of results within most of the program categories.

To further analyze these results, we conducted several sensitivity analyses to test for confounds of the significant differences we found between program types at postintervention. These analyses revealed that (a) PCIT had a higher percentage of clinical applications (all 14 comparisons were either indicated preventions or clinical applications) that generally showed higher mean effect sizes (see below), and (b) that effect sizes based more on observational data (7 of 14 comparisons) showed higher mean effect sizes as well (see above). In addition, and most importantly, PCIT comparisons had much lower (about one *SD*) mean sample sizes ($M = 58.2$, $k = 12$ for PFO and $M = 51.3$, $k = 14$ for ASB) compared to the mean sample size of all comparisons ($M = 94.7$, $k = 226$ for PFO and $M = 96.32$, $k = 258$ for ASB) that, again, had some positive impact on effect sizes (see below). The same holds even more for humanistic programs ($M = 45.1$, $k = 10$ for PFO and $M = 42.5$, $k = 12$ for ASB) and explained at least in part the slightly higher mean effect sizes displayed in Table 4.

When we compared program types on the most global level between cognitive-behavioral programs and (the heterogeneous group of) nonbehavioral programs and mixed programs, we found some significant differences in PFO favoring nonbehavioral programs ($d_+ = 0.44$, $k = 196$, $d_+ = 0.72$, $k = 22$, and $d_+ = 0.38$, $k = 8$, respectively), Q_b ($df = 2$) $= 10.57$, $p < .01$, and comparable results on ASB ($d_+ = 0.46$, $k = 222$, $d_+ = 0.58$, $k = 26$, and $d_+ = 0.40$, $k = 10$, respectively), Q_b ($df = 2$) $= 2.19$, $p < .34$. In addition, follow-up results on program types revealed the same picture, but several program types had only a few follow-up studies. However, some results are worth mentioning: Four program types had more than two comparisons at short-term follow up (3 to 12 months after the termination of the program): Triple P had significant effects on PFO ($d_+ = 0.36$, $k = 10$) and ASB ($d_+ = 0.43$, $k = 10$). Incredible Years also had significant but somewhat lower effect sizes ($d_+ = 0.29$, $k = 11$ on PFO, and $d_+ = 0.27$, $k = 12$ on ASB). Parent Management Training had low effects on PFO ($d_+ = 0.20$, $k = 3$) and nonsignificant results on ASB ($d_+ = 0.03$, $k = 6$) and other cognitive-

Table 4
Postintervention effect sizes according to program type, level of intervention, and age of children/adolescents.

Characteristic	Parent/family outcomes				Antisocial behavior outcomes			
	$d_+ \pm$ CI	k	Q_w	Q_b (df)	$d_+ \pm$ CI	k	Q_w	Q_b (df)
Program name/type/approach				30.04 (8)***				17.10 (8)*
Triple P	0.39 ± 0.09	51	20.12		0.44 ± 0.10	55	59.83	
Incredible Years	0.46 ± 0.12	34	26.98		0.42 ± 0.12	37	19.28	
Parent Management Training OM	0.46 ± 0.18	12	5.37		0.35 ± 0.17	16	12.16	
Parent Child Interaction Therapy	0.87 ± 0.22	12	21.83*		0.83 ± 0.22	14	20.11+	
Other cognitive-behavioral programs	0.40 ± 0.07	87	121.71*		0.46 ± 0.07	100	101.33	
Adlerian programs	0.73 ± 0.27	8	13.97+		0.61 ± 0.26	10	14.50	
Humanistic programs	0.82 ± 0.26	10	26.37**		0.66 ± 0.25	12	15.55	
Attachment/relationship-based programs	0.53 ± 0.33	4	2.89		0.35 ± 0.34	4	5.77	
Mixed programs	0.37 ± 0.22	8	11.83*		0.40 ± 0.22	10	31.62***	
Intervention level				7.97 (3)*				10.38 (3)**
Universal prevention	0.42 ± 0.09	57	50.72		0.38 ± 0.08	64	53.85	
Selective prevention	0.45 ± 0.10	59	106.82***		0.41 ± 0.09	66	89.40*	
Indicated prevention	0.43 ± 0.09	68	68.27		0.51 ± 0.08	80	108.42*	
Clinical applications	0.62 ± 0.12	42	32.92		0.59 ± 0.11	48	41.08	
Age of child/adolescent				5.27 (3)				7.95 (3)*
Up to 4 years	0.45 ± 0.11	49	60.78		0.46 ± 0.10	51	61.92	
4–7 years	0.51 ± 0.07	100	113.90		0.54 ± 0.07	118	135.91	
8–10 years	0.47 ± 0.11	40	43.69		0.39 ± 0.10	49	58.47	
11–18 years	0.35 ± 0.11	37	39.49		0.38 ± 0.12	40	33.64	

Notes. OM = Oregon Model; d_+ = weighted mean effect size; CI = limits of the 95 % confidence interval; k = number of comparisons; Q_w = homogeneity statistic within categories; Q_b = homogeneity statistic between categories.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

behavioral programs with significant results ($d_+ = 0.25$, $k = 22$ on PFO, and $d_+ = 0.19$, $k = 24$ on ASB). Only three program types investigated long-term effects 12 months or more after the termination of the program with more than two comparisons. For Triple P programs, we found a significant effect size of $d_+ = 0.25$, $k = 6$ on PFO, but a nonsignificant effect size for ASB ($d_+ = 0.06$, $k = 7$). Incredible Years programs yielded a significant effect size on PFO ($d_+ = 0.17$, $k = 11$) and a nonsignificant effect size on ASB ($d_+ = 0.10$, $k = 11$). Finally, cognitive-behavioral programs had a nonsignificant effect on PFO ($d_+ = 0.06$, $k = 8$) and a small but significant effect on ASB ($d_+ = 0.15$, $k = 13$).

Table 4 also show the effects on several levels of interventions. Overall, we found significantly higher effects on PFO and ASB for clinical applications. Otherwise, there were surprisingly only small differences between the various types of prevention levels, although we obtained an increase in effect sizes according to the risk level of the target group in ASB outcomes. Again, as in the analyses of program type, we found that the majority of categories showing within homogeneity displayed relatively small effect-size variation and thus no indication for further moderators. However, we did find larger differences between different levels of interventions in the follow-up data. Although effect sizes were relatively equal for PFO at short- and long-term follow up (effect sizes ranged from $d_+ = 0.16$ to 0.37 between the different levels of intervention at short-term, and from $d_+ = 0.10$ to 0.16 at long-term follow up), there were marked effect-size differences on ASB outcomes (Q_b [$df = 3$] = 17,66, $p < .001$ for short-term follow up, and Q_b [$df = 3$] = 7.47, $p < .06$ for long-term follow up). Whereas universal and selective prevention had very low effects in the 3- to12-month follow-up period ($d_+ = 0.14$, $k = 12$, $p < .05$, and $d_+ = 0.06$, $k = 17$, respectively), we found significant effects for indicated prevention and clinical applications of PTP ($d_+ = 0.34$, $k = 22$, and $d_+ = 0.47$, $k = 7$, respectively). On a lower level, the same pattern was observed in the long-term follow-up data (e.g., 12 months or more: $d_+ = 0.05$, $k = 9$ for universal, and $d_+ = 0.07$, $k = 14$ for selective prevention; $d_+ = 0.29$, $k = 9$, $p < .01$ for indicated prevention, and $d_+ = 0.22$, $k = 2$ clinical applications), although the data base, especially for clinical applications, was rather small.

In addition, age of children/adolescents was not a strong moderator of PTP outcomes, although we found a slight negative tendency for PFO ($\beta = -0.14$, $k = 218$, $p < .05$) but not for ASB outcomes ($\beta = -0.09$, $k = 249$) at postintervention. Results for different age groups are displayed in Table 4. Effect sizes indicated no significant differences between age groups on PFO and low but significant differences for ASB favoring the 4- to 7-year-old age group. This group was also the main target group and had the best follow-up results (significant effects at short- and long-term follow up for ASB, $d_+ = 0.34$, $k = 21$, and $d_+ = 0.19$, $k = 14$, respectively). Again, the age categories showed highly homogeneous results.

4.6. Further effect-size moderators

As reported above, we found a clear negative linear trend between sample and effect sizes that indicated some degree of publication bias. However, in intervention research, this negative connection can also be explained by a better implementation in smaller studies that, in turn, leads to higher effect sizes (Beelmann et al., 2018). Therefore, we categorized sample size into four groups (up to 30, 31 to 50, 51 to 100, and above 100), and found clear outcome differences between these groups at postintervention. Moreover, there was a continuous decrease in effect size for ASB outcomes ($d_+ = 0.78$, $k = 44$; $d_+ = 0.65$ $k = 64$; $d_+ = 0.50$, $k = 82$; $d_+ = 0.26$; $k = 67$, respectively; all results were significant and homogeneous) leading to highly significant differences, Q_b ($df = 3$) = 72.46, $p < .001$. Almost the same could be obtained for PFO ($d_+ = 0.71$, $k = 35$; $d_+ = 0.63$ $k = 52$; $d_+ = 0.46$, $k = 76$; $d_+ = 0.34$; $k = 60$, respectively; Q_b [$df = 3$] = 28.48, $p < .001$). A final analysis addressed the independence of the evaluator, because prior analyses (Eisner & Humphreys, 2012) have found this to be an important moderator. However, when we compared a group of comparisons with a high dependency between program developer and evaluator (e.g., same person or author or co-author) with comparisons with medium and low dependency (e.g., foreign evaluator investigates an international program, no dependency between developer and evaluator), we found some nonsignificant differences ($d_+ = 0.50$, $k = 156$, and $d_+ = 0.45$, $k = 92$, respectively, for ASB; and $d_+ = 0.48$, $k = 137$, and $d_+ = 0.45$, $k = 80$, respectively, for PFO).

5. Discussion

This meta-analysis integrating 239 reports, 241 studies, and 279 comparisons between an intervention and a control group with >2800 effect sizes clearly confirms the general finding of prior reviews: Parent training programs have substantial short-term effects on antisocial behavior in children and adolescents (see, e.g., Beelmann & Raabe, 2009; Farrington, Gaffney, Lösel, et al., 2017; Mingebach et al., 2018). The statistical material shows a significant medium effect size on a broad range of parent and family outcomes ($d = 0.46$) as well as on measures of antisocial behavior ($d = 0.47$) at postintervention—that is, up to 3 months after the termination of the program. These results hold for a great variety of programs, target groups, and ages. Compared to other parent training meta-analyses, these findings are based on a larger number of primary studies providing a detailed analysis of different outcome assessments of parent and family as well as ASB measures and including a full range of programs, ages, and levels of interventions (Weber et al., 2019).

Moreover, we checked how changes in parental skills and other parental and family variables are linked to changes in ASB outcomes, thus testing a general assumption of PTP. These mediational analyses reveal a clear link between parent and family outcomes and reductions in antisocial behavior as Fagan and Benedini (2016) were able to confirm in a narrative analysis. Indeed, except for more distal outcomes such as marital satisfaction, measures of family quality (cohesion, few conflicts), or parent psychopathology, changes in parental skills and characteristics show considerable links to changes in child or adolescent antisocial behavior outcomes. The best predictors of ASB outcomes are reductions in parental stress followed by the promotion of parent–child interaction/relation, reductions in aversive parental behavior, and the strengthening of parental competencies (e.g., praise, reward, supervision, effective limit setting). This is also in line with longitudinal developmental reviews (e.g., Flanagan et al., 2019; Pinquart, 2017).

A closer look at different ASB outcomes reveals a few differences showing that outcomes on general ASB rating scales and on oppositional/disruptive behavior are somewhat higher with small follow-up effects. In contrast, data show that it is particularly effect sizes for aggressive and delinquent behavior that are slightly lower at postintervention and nonsignificant at follow up. In addition, only a few findings are based on "hard data" such as school and police records, and most outcomes were measured using general ASB rating scales. However, because the main target group is parents with 3- to 10-year-old children, this limitation seems understandable.

In addition to the overall findings, analyses of different program types and approaches reveal hardly any reliable differences. Although brand-name programs such as the PCIT and humanistic approaches display higher effect sizes on PFO and ASB, we found that their results are confounded with the level of intervention (higher effect sizes for clinical applications), the outcome assessment (higher effect sizes for observational data), and the sample size (higher effect sizes in smaller studies). Thus, all analyzed programs and program types and approaches produce just about the same evidence for changing parent and family-related outcomes and antisocial behavior in children and adolescents. Furthermore, as most of the analyzed categories yield homogeneous results, we could not expect further moderators. Because the content of the diverse programs shows clear overlaps, we naturally cannot expect too many differences in effectiveness. This is especially the case for the

great number of cognitive-behavioral programs, but also for similarities between these and nonbehavioral programs. For example, the STEP program (Adlerian tradition) and the PET program (humanistic tradition) also include the learning of behavioral management strategies (see, www.steppublishers.com; www.gordontraining.com). Inversely, many cognitive-behavioral programs also include exercises on parent–child communication and are directed toward the promotion of parent–child relations (see, e.g., Leijten et al., 2018). This makes it difficult if not impossible to compare different programs or to study program components independently via meta-analytic data.

When we compared the evidence on PTP with other prevention or intervention alternatives, we found relatively high effects at post-intervention. For example, child social skills trainings are also used very frequently in the prevention of ASB alone and in combination with PTP. A recent meta-analysis (Beelmann & Lösel, 2021) has confirmed significant effects on ASB but with a lower mean effect size ($d = 0.25$). Missing or at least decreasing long-term effects are also an issue with these programs. Even more intensive approaches such as early intervention or early education programs, which have been used mostly to promote the development of young children at risk, fail to be more successful. Some of these projects have been studied in extended long-term evaluations that also found low or non-significant effect sizes (see Deković et al., 2011). In addition, the most intensive and combined approaches delivered within a cross-linked system of help such as the Community That Cares (CTC) approach recently found no long-term effects on ASB and crime on a community level (Fagan et al., 2019; Rhew et al., 2016). Therefore, there is still a lack of evidence for convincing and practically significant long-term effects—especially for low- and medium-intensity prevention measures in the field of ASB and crime. However, a low mean effect size within variable-centered analyses does not mean that the programs have no practical significance. Some projects show that they lead to fewer problem cases or higher proportions of healthy children and adolescents in long-term follow-up studies (see, e.g., the results of the Fast Track Program; Bierman et al., 2020 and Dodge et al., 2015; or Reynolds et al., 2011; Schweinhart, 2013). Such person-centered analyses should be an issue for future research on PTP as well as on other low-intensity prevention programs. For example, one long-term evaluation of the German Triple P-Program with a 10-year follow up showed that PTP can have at least small long-term effects by reducing the incidence of problematic cases with externalizing behavior (Hahlweg & Schulz, 2018).

Interestingly, we also found relatively small differences due to the level of intervention. To some extent, this runs counter to findings from other recent reviews (Leijten et al., 2019). It may have to do with different data sets (e.g., different age ranges) or with the way we calculated effect sizes (e.g., we integrated effect sizes within comparisons to avoid dependency). However, like other authors, we confirmed that clinical applications of PTP lead to higher mean effect sizes (e.g., de Graaf, Speetjens, Smit, de Wolff, Tavecchio, 2008a, 2008b; Leijten et al., 2019; Lundahl et al., 2006; Menting et al., 2013), but we did not find big differences between the various preventive levels. This is especially noteworthy for universal applications, because they usually lead to lower effects because of the lower base rate of either problem behavior and bottom effects or higher competence scores and ceiling effects (see Beelmann et al., 2018; Sandler et al., 2014). However, conditions could be different in universal PTP, especially for parents of younger children, because they acquire new parental competencies that had not been developed before participation in the program. In addition, because ASB is defined as deviation from age-appropriate social behavior, it naturally becomes a stronger problem from early preschool age on and generally has a high base rate over the course of development. Thus, a universal prevention application of PTP may have the same or nearly the same potential to change parent and child competencies and behavior as selective or indicative preventions. In addition, short- and long-term follow-up effects show the expected pattern in which indicated prevention and clinical applications display higher effect sizes than universal or selective preventions, although decreasing effect sizes are obtained on all levels.

The age of the children or adolescents is only a minor moderator of the impact of PTP. We found a small linear trend in favor of younger children and also better results for these age groups at follow up. Other researchers also found a low impact of age on the effectiveness of PTP (Gardner et al., 2019), but results may be restricted because most programs are directed at 3- to10-year-olds. Some authors have discussed these findings in the light of questioning the preventive idea of "earlier is better." However, systematic comparisons of early and late application of PTP are necessary to address this issue more adequately. In addition, developmentally informed prevention models do not suggest that earlier is better in general, but that programs should be timely and age-appropriate or, better, developmentally appropriate (Beelmann, 2012; Beelmann et al., 2018; Malti et al., 2016). At the moment, we can say that the use of PTP is effective in all age groups, although we have to assume that the impact of parents on their children decreases over the course of development. But the positive effects of PTP on older age groups (e.g., adolescents) may indicate that there are important functions of parenting behavior at these ages as well (e.g., delivery of opportunities for identity development) that probably differ from the functions in younger children (support in autonomy, limit setting for inappropriate behavior).

Despite the quite positive overall results for PTP evaluations, we identified several findings that dampen the general positive picture. First, we found some indication of publication bias that probably reduces the mean effect size. However, an alternative interpretation would be that small studies enable a better implementation and thus lead to higher effect sizes in studies with lower sample sizes. We coded several implementation measures (e.g., attendance rate, satisfaction with the program, etc.), but found no reliable implementation index that had been measured across any significant number of studies. Therefore, whether the high negative correlation between sample size and effect size indicates publication bias or is a result of a higher implementation quality in smaller studies remains an open question. Whatever the case, sample size seems to be the strongest moderator of effectiveness and even stronger than program type or level of intervention—as also found in other meta-analyses (e.g., Beelmann & Lösel, 2021). Therefore, in this situation, it seems advisable to focus less on program components or characteristics of the application (level, age), but more on two other aspects: first, a thorough unified measurement of implementation quality of PTP. In recent decades, this has been suggested by several authors in intervention science in general (e.g., Ghate, 2016) and in parenting interventions in particular (Mauricio et al., 2018; Weisenmuller & Hilton, 2020). Second, it is necessary to ensure high implementation quality by delivering high-standard implementation systems (e.g., Meyers et al., 2012) that guarantee service delivery with fidelity and situational flexibility including culturally adapted concepts (Baumann et al., 2015; Beelmann et al., 2021; Kumpfer et al., 2017; Sundell et al., 2016).

A second caveat for the effects of PTP is the continuously decreasing follow-up effects from about $d = 0.47$ to 0.22 and 0.12 for ASB and from $d = 0.46$ to 0.27 and 0.13 for PFO. This reduction is found for nearly all programs and all applications and has already been addressed by Leijten et al. (2019). However, our results do not confirm other review analyses (e.g., van Aar et al., 2017) that found stable follow-up effects. But in contrast to other reviews, our analysis is based on a broader data base and unified follow-up periods (e.g., 3 to 12 months, 12 months or more) that are independent from the follow-up definition in the primary studies, thereby allowing us to summarize comparable follow-up effects. Thus, although we found small but significant follow-up effect sizes in many cases, our findings clearly reveal a reduced stability of PTP impacts on parents and families as well as on antisocial behavior outcomes. One way to stabilize effects could be a repeated age-appropriate application of PTP or—as mentioned above—an imbedded implementation and delivery within a broader concept of child and family services.

Another limitation to the findings concerns the restricted outcome assessment despite the great variation in methods used. A large majority of effect sizes are based on ratings by parents that probably suffer from reactivity and social desirability. Although we found higher effect sizes for observational methods using independent raters, these data are mostly collected by special observational tests such as the DPICS and not by natural observational assessments. In contrast, the few ratings by teachers reveal no significant effect on ASB. Thus, it seems advisable in the future to test effects with more emphasis on their generalizability across different informants and settings. In this regard, it is also interesting that the main PFO predictor of ASB changes is perceived parental stress (see above). This may indicate that parents—at least to some extent—simply adopt a different view of the problem behavior of their children that, in itself, has not changed too much but is simply being rated as less disturbing or problematic.

Finally, we found some indications that the dependence on evaluator/study author and program developer has some impact on the results. However, these differences are not as pronounced as to be expected from prior results (Eisner & Humphreys, 2012). Nonetheless, we have to take some conflict of interests into account, because the majority of studies stem from the developer of the programs or related research groups (Gorman, 2018).

6. Conclusions

Although there have been numerous applications of PTP in the prevention and treatment of antisocial behavior going back over many years, we still find open questions (how to explain the connection between sample and effect sizes) and scope for improvements (e.g., long-term effects). Nonetheless, we should not underestimate the short-term benefits of parent training programs on a number of parental competencies and child behavioral development—benefits that have been found in a series of high-quality evaluation studies around the world. We cannot expect that small and less intensive programs can lead to sustained effects in every single case or protect all children and adolescents from a deviant development, but we should nevertheless improve promising programs if we are to make things better for parents and their children.

Funding

This review was supported by a Grand of the Jacobs Foundation, Zurich, Switzerland.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material (data transparency)

Not applicable.

Code availability (software application or custom code)

Not applicable.

Submission declaration

The manuscript is not under consideration for publication elsewhere.

Declaration of competing interest

The authors report no conflicts of interest or competing interests.

Data availability

No data was used for the research described in the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.avb.2022.101798.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Association.

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Barlow, J., Parsons, J., & Stewart-Brown, S. (2005). Preventing emotional and behavioural problems: The effectiveness of parenting programmes with children less than 3 years of age. *Child: Care, Health, and Development, 31*(1), 33–42. https://doi.org/10.1111/j.1365-2214.2005.00447.x

Baumann, A. A., Powell, B. J., Kohl, P. L., Tabak, R. G., Penalba, V., Proctor, E. K., Domenech-Rodriguez, M. M., & Cabassa, L. J. (2015). Cultural adaptation and implementation of evidence-based parent-training: A systematic review and critique of guiding evidence. *Children and Youth Services Review, 53*, 113–120. https://doi.org/10.1016/j.childyouth.2015.03.025

Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development, 37*(4), 887–907. https://doi.org/10.2307/1126611

Beelmann, A. (2012). The scientific foundation of prevention. The status quo and future challenges of developmental crime prevention. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime. Contributions of developmental and evaluation research to prevention and intervention* (pp. 137–164). Hogrefe.

Beelmann, A., Arnold, L. S., & Schulz, S. (2021). Buffering negative effects of immigration on cognitive, social, and educational development: A multinational meta-analysis of child and adolescent prevention programs. *International Journal of Psychology, 56*(3), 478–490. https://doi.org/10.1002/ijop.12725

Beelmann, A., & Lösel, F. (2021). A comprehensive meta-analysis of randomized evaluations of the effect of child skills training on antisocial development. *Journal of Developmental and Life-Course Criminology, 7*(1), 41–65. https://doi.org/10.1007/s40865-020-00142-8

Beelmann, A., Malti, T., Noam, G., & Sommer, S. (2018). Innovation and integrity: Desiderata and future directions for prevention and intervention science. *Prevention Science, 19*(3), 358–365. https://doi.org/10.1007/s11121-018-0869-6

Beelmann, A., & Raabe, T. (2009). The effects of preventing antisocial behavior and crime in childhood and adolescence. Results and implications of research reviews and meta-analyses. *International Journal of Developmental Science, 3*(3), 260–281. https://doi.org/10.3233/DEV-2009-3305

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088–1101. https://doi.org/10.2307/2533446

Belfer, M. L. (2008). Child and adolescent mental disorders: The magnitude of the problem across the globe. *Journal of Child Psychology and Psychiatry, 49*(3), 226–236. https://doi.org/10.1111/j.1469-7610.2007.01855.x

Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. E. (2020). *The Fast Track Program for children at risk. Preventing antisocial behavior. Guilford*.

Bowlby, J. (1969). *Attachment and loss. Vol. I: Attachment*. Basic Books.

Bowlby, J. (1973). *Attachment and loss. Vol. II: Separation*. Basic Books.

Braga, T., Gonçalves, L. C., Basto-Pereira, M., & Maia, Â. (2017). Unraveling the link between maltreatment and juvenile antisocial behavior: A meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior, 33*, 37–50. https://doi.org/10.1016/j.avb.2017.01.006

Briesmeister, J. M., & Schaeffer, C. M. (2007). *Handbook of parent training. Parents as co-therapists for children's behavior problems* (3rd ed.). Wiley.

Burke, J. D., Loeber, R., & Birmaher, B. (2002). Oppositional defiant disorder and conduct disorder: A review of the past 10 years, part II. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(11), 1275–1293. https://doi.org/10.1097/00004583-200211000-00009

Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of parent effectiveness training. *The American Journal of Family Therapy, 18*(4), 373–384. https://doi.org/10.1080/01926189008250986

de Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008a). Effectiveness of the triple P positive parenting program on parenting: A meta-analysis. *Family Relations, 57*(5), 553–566. https://doi.org/10.1111/j.1741-3729.2008.00522.x

de Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of the Triple P Positive Parenting Program on behavioral problems in children. A meta-analysis. *Behavior Modifications, 32*(5), 714–735. https://doi.org/10.1177/0145445508317134

Deković, M., Slagt, M. I., Asscher, J. J., Boendermaker, L., Eichelsheim, V. I., & Prinzie, P. (2011). Effects of early prevention programs on adult criminal offending: A meta-analysis. *Clinical Psychology Review, 31*(4), 532–544. https://doi.org/10.1016/j.cpr.2010.12.003

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *JAMA, 263*(10), 1385–1389. https://doi.org/10.1001/jama.1990.03440100097014

Dodge, K. A., Bierman, K. L., Coie, J. D., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. E. (2015). Impact of early intervention on psychopathology, crime, and well-being at age 25. *American Journal of Psychiatry, 172*(1), 59–70. https://doi.org/10.1176/appi.ajp.2014.130600786

Dreikurs, R., & Soltz, V. (1964). *Children: The challenge*. Hawthorn Books.

Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., Taylor, R. S., Sandercock, J., & Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomized controlled trials. *Child and Adolescent Psychiatry and Mental Health, 3*(1), 7. https://doi.org/10.1186/1753-2000-3-7

Dretzke, J., Frew, E., Davenport, C., Barlow, J., Stewart-Brown, S., Sandercock, J., Bayliss, S., Raftery, J., Hyde, C., & Taylor, R. (2005). The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorders, including oppositional defiant disorder, in children. *Health Technology Assessment, 9*(50), 1–233. https://doi.org/10.3310/hta9500

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Eisner, M., & Humphreys, D. (2012). Measuring conflict of interest in prevention and intervention research. A feasibility study. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime. Contributions of developmental and evaluation research to prevention and intervention* (pp. 165–180). Hogrefe.

Eyberg, S. M., & Boggs, S. R. (1998). Parent-child interaction therapy: A psychosocial intervention for the treatment of young conduct-disordered children. In J. M. Briesmeister, & C. E. Schaefer (Eds.), *Handbook of parent training. Parents as co-therapists for children's behavior problems* (pp. 61–97). Wiley.

Fagan, A., Hawkins, J., Catalano, R., & Farrington, D. (2019). Findings from the community youth developmental study. In A. Fagan, J. Hawkins, D. Farrington, & J. Catalano (Eds.), *Communities that care: Building community engagement and capacity to prevent youth behavior problem* (pp. 179–210). Oxford University Press.

Fagan, A. A., & Benedini, K. M. (2016). How do family-focused prevention programs work? A review of mediating mechanism associated with reductions in youth antisocial behaviors. *Clinical Child and Family Psychology Review, 19*(4), 285–309. https://doi.org/10.1007/s10567-016-0207-0

Farrington, D. P. (2007). Childhood risk factors and risk-focused prevention. In M. Meguire, R. Morgan, & R. Reiner (Eds.), *Vol. 4. The Oxford handbook of criminology* (pp. 602–640). Oxford University Press.

Farrington, D. P., Gaffney, H., Lösel, F., & Ttofi, M. (2017). Systematic reviews of the effectiveness of developmental prevention programs in reducing delinquency, aggression, and bullying. *Aggression and Violent Behavior, 33*, 91–106. https://doi.org/10.1016/j.avb.2016.11.003

Farrington, D. P., Gaffney, H., & Ttofi, M. M. (2017). Systematic reviews of explanatory risk factors for violence, offending, and delinquency. *Aggression and Violent Behavior, 33*, 24–36. https://doi.org/10.1016/j.avb.2016.11.004

Farrington, D. P., Loeber, R., & Ttofi, M. M. (2012). Risk and protective factors for offending. In B. C. Welsh, & D. P. Farrington (Eds.), *The Oxford handbook of crime prevention* (pp. 46–69). Oxford University Press.

Farrington, D. P., & Welsh, B. C. (2003). Family-based prevention of offending: A meta-analysis. *Australian & New Zealand Journal of Criminology, 36*(2), 127–151. https://doi.org/10.1375/acri.36.2.127

Fergusson, D. M., Horwood, L. J., & Ridder, E. M. (2005). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry, 46*(8), 837–849. https://doi.org/10.1111/j.1469-7610.2004.00387.x

Flanagan, I. M. L., Auty, K. M., & Farrington, D. P. (2019). Parental supervision and later offending: A systematic review of longitudinal studies. *Aggression and Violent Behavior, 47*, 215–229. https://doi.org/10.1016/j.avb.2019.06.003

Forehand, R., Jones, D. J., & Parent, J. (2013). Behavioral parenting interventions for child disruptive behaviors and anxiety: what's different and what's the same. *Clinical Psychology Review, 33*(1), 133–145. https://doi.org/10.1016/j.cpr.2012.10.010

Forehand, R., Lafko, N., Parent, J., & Burt, K. B. (2014). Is parenting the mediator of change in behavioral parent training for externalizing problems of youth? *Clinical Psychology Review, 34*(8), 608–619. https://doi.org/10.1016/j.cpr.2014.10.001

Frick, P. J. (1998). *Conduct disorders and severe antisocial behavior*. Plenum.

Frick, P. J., Lahey, B. B., Loeber, R., Tannenbaum, L., van Horn, Y., Christ, M. A. G., Hart, E. A., & Hanson, K. (1993). Oppositional defiant disorder and conduct disorder: A meta-analytic review of factor analyses and cross-validation in a clinic sample. *Clinical Psychology Review, 13*(4), 319–340. https://doi.org/10.1016/0272-7358(93)90016-F

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smit, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Campbell Systematic Reviews, 8*(1), 1–239. https://doi.org/10.4073/csr.2012.12

Gardner, F., Leitjen, P., Melendez-Torres, G. J., Landau, S., Harris, V., Mann, J., Beecham, J., Hutching, J., & Scott, S. (2019). The earlier the better? Individual participant data and traditional meta-analysis of age effects on parenting interventions. *Child Development, 90*(1), 7–19. https://doi.org/10.1111/cdev.13138

Ghate, D. (2016). From programs to systems: Deploying implementation science and practice for sustained real world effectiveness in services for children and families. *Journal of Clinical Child and Adolescent Psychology, 45*(6), 812–826. https://doi.org/10.1080/15374416.2015.1077449

Gilbody, S. M., Song, F., Eastwood, A. J., & Sutton, A. (2000). The causes, consequences and detection of publication bias in psychiatry. *Acta Psychiatrica Scandinavica, 102*(4), 241–249. https://doi.org/10.1034/j.1600-0447.2000.102004241.x

Gordon, T. (1970). *Parent effectiveness training*. Wyden.

Gorman, D. M. (2018). Can we trust positive findings of intervention research? The role of conflict of interest. *Prevention Science, 19*(3), 295–305. https://doi.org/10.1007/s11121-016-0648-1

Granic, I., & Patterson, G. R. (2006). Toward a comprehensive model of antisocial development: A dynamic systems approach. *Psychological Review, 113*(1), 101–131. https://doi.org/10.1037/0033-295X.113.1.101

Hahlweg, K., & Schulz, W. (2018). Universelle Prävention kindlicher verhaltensstörungen durch elterntraining: wirksamkeit nach 10 jahren aus sicht der mütter, väter und jugendlichen [universal prevention of childhood behavioral disorders by parent training: 10-year effectiveness from mothers', fathers', and adolescents' perspectives]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 47*(1), 1–15. https://doi.org/10.1026/1616-3443/a000462

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Academic Press.

Hoeve, M., Dubas, J. S., Eichelsheim, V. I., van der Laan, P. H., Smeenk, W., & Gerris, J. R. M. (2009). The relationship between parenting and delinquency: A meta-analysis. *Journal of Abnormal Child Psychology, 37*(6), 749–775. https://doi.org/10.1007/s10802-009-9310-8

Jolliffe, D., Farrington, D. P., Piquero, A. R., Loeber, R., & Hill, K. G. (2017). Systematic review of early risk factors for life-course-persistent, adolescence-limited, and late-onset offenders in prospective longitudinal studies. *Aggression and Violent Behavior, 33*, 15–23. https://doi.org/10.1016/j.avb.2017.01.009

Kazdin, A. E. (2005). *Parent management training. Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents*. Oxford University Press.

Knerr, W., Gardner, F., & Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prevention Science, 14*(4), 352–363. https://doi.org/10.1007/s11121-012-0314-1

Kumpfer, K., Magalhães, C., & Xie, J. (2017). Cultural adaptation and implementation of family evidence-based interventions with diverse population. *Prevention Science, 18*(6), 649–659. https://doi.org/10.1007/s11121-016-0719-3

Lahey, B. B., Miller, T. L., Gordon, R. A., & Riley, A. W. (1999). Developmental epidemiology of the disruptive behavior disorders. In H. C. Quay, & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 23–48). Kluwer Academic / Plenum.

Leijten, P., Gardner, F., Melendez-Torres, G. J., van Aar, J., Hutchings, J., Schulz, S., Knerr, W., & Overbeek, G. (2019). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the Academy of Child & Adolescent Psychiatry, 58*(2), 180–190. https://doi.org/10.1016/j.jaac.2018.07.900

Leijten, P., Melendez-Torres, Gardner, F., van Aar, J., Schulz, S., & Overbeek, G. (2018). Are relationship enhancement and behavior management „The Golden Couple" for disruptive child behavior? Two meta-analyses. *Child Development, 89*(6), 1970–1982. https://doi.org/10.1111/cdev.13051

Leijten, P., Raijmakers, M. A. J., 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

Lipsey, M. W., & Derzon, J. H. (1998). Predictors of violent and serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber, & D. P. Farrington (Eds.), *Serious & violent juvenile offenders. Risk factors and successful interventions* (pp. 86–105). Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Sage.

LoBraico, E. J., Bray, B. C., Feinberg, M. E., & Fosco, G. M. (2020). Constellations of family risk for long-term adolescent antisocial behavior. *Journal of Family Psychology, 34*(5), 587–597. https://doi.org/10.1037/fam0000640

Loeber, R., Burke, J. D., Lahey, B. B., Winters, A., & Zera, M. (2000). Oppositional defiant and conduct disorder: A review of the past 10 years, part I. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(12), 1468–1484. https://doi.org/10.1097/00004583-200012000-00007

Lösel, F., & Bender, D. (2016). Parenting and family-oriented programmes for parenting child behavior problems: What the evidence tells us. *Journal of Family Research, 11*, 217–239. https://doi.org/10.3224/zff.s11y2016/2017.12

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. https://doi.org/10.1016/j.cpr.2005.07.004

Lundahl, B. W., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice, 18*(2), 97–106. https://doi.org/10.1177/1049731507309828

Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine, 20*(4), 641–654. https://doi.org/10.1002/sim.698

Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interaction (Vol. Ed.). In P. H. Mussen (Ser. Ed.), & E. M. Hetherington (Eds.) (4th ed.,*Socialization, personality, and social development: Vol. 4. Handbook of child psychology* (pp. 1–101). Wiley.

Malti, T., Noam, G. G., Beelmann, A., & Sommer, S. (2016). Toward dynamic adaptation of psychological interventions for child and adolescent development and mental

health. *Journal of Clinical Child and Adolescent Psychology, 45*(6), 827–836. https://doi.org/10.1080/15374416.2016.1239539
Matjasko, J. L., Vivolo, A. M., Massetti, G. M., Holland, K. M., Holt, M. K., & Cruz, J. D. (2012). A systematic meta-review of evaluations of youth violence prevention programs: Common and divergent findings from 25 years of meta-analyses and systematic reviews. *Aggression and Violent Behavior, 17*, 540–552. https://doi.org/10.1016/j.avb.2012.06.006
Maughan, B., Rowe, R., Messer, J., Goodman, R., & Meltzer, H. (2004). Conduct disorder and oppositional defiant disorder in a national sample: Developmental epidemiology. *Journal of Child Psychology and Psychiatry, 45*(3), 609–621.
Maughan, D., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review, 34*(3), 267–286. https://doi.org/10.1080/02796015.2005.12086287
Mauricio, A. M., Gonzales, N. A., & Sandler, I. N. (2018). Preventive parenting interventions: Advancing conceptualizations of participation and enhancing reach. *Prevention Science, 19*(5), 603–608. https://doi.org/10.1007/s11121-018-0876-7
McCard, M. R., Priester, P. E., Davis, W. H., & Azen, R. (2006). Differential effectiveness of behavioural parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology, 34*(4), 527–543. https://doi.org/10.1007/s10802-006-9031-1
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), 901–913. https://doi.org/10.1016/j.cpr.2013.07.006
Meyers, D. C., Durlak, J. A., & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology, 50*(3–4), 462–480. https://doi.org/10.1007/s10464-012-9522-x
Mingebach, T., Kamp-Becker, I., Christiansen, H., & Weber, L. (2018). Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. *PLOS ONE, 13*(9), Article e0202855. https://doi.org/10.1371/journal.pone.0202855
Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674–701. https://doi.org/10.1037/0033-295X.100.4.674
Moore, A. A., Silberg, J. L., Roberson-Nay, R., & Mezuk, B. (2017). Life course persistent and adolescence limited conduct disorder in a nationally representative US sample: Prevalence, predictors, and outcomes. *Social Psychiatry and Psychiatric Epidemiology, 52*(4), 435–443. https://doi.org/10.1007/s00127-017-1337-5
Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386, 10.1177/1094428106291059.
Muñoz, R. F., Hutchings, J., Edwards, R. T., Hounsome, B., & O'Ceilleachair, A. (2004). Economic evaluation of treatments for children with severe behavioural problems. *Journal of Mental Health Policy Economics, 7*(4), 177–189.
National Institute for Health and Care Excellence. (2013, March 27). Antisocial behaviour and conduct disorders in children and young people: Recognition and management. https://www.nice.org.uk/guidance/cg158.
Newcombe, R. G. (1987). Towards a reduction in publication bias. *British Medical Journal (Clinical Research Ed.), 295*(6599), 656–659. https://doi.org/10.1136/bmj.295.6599.656
Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of triple P-positive parenting program using hierarchical linear modelling: Effectiveness and moderating variables. *Clinical Child and Family Psychology Review, 11*(3), 114–144. https://doi.org/10.1007/s10567-008-0033-0
Patterson, G. R. (1982). *Coercive family process*. Castalia.
Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys: A social interactional approach* (vol. 4). Castalia.
Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology, 53*(5), 873–932. https://doi.org/10.1037/dev00000295
Piquero, A., Farrington, D. P., Welsh, B., Tremblay, R., & Jennings, W. (2008). Effects of early family/parent training programs on antisocial behaviour & delinquency. A systematic review. *Campbell Systematic Reviews, 4*(1), 1–122. https://doi.org/10.4073/csr.2008.11
Piquero, A., Farrington, D. P., Welsh, B., Tremblay, R., & Jennings, W. (2009). Effects of early family/parent training programs on antisocial behaviour & delinquency. *Journal of Experimental Criminology, 5*(2), 83–120. https://doi.org/10.1007/s11292-009-9072-x
Piquero, A. R., Jennings, W. G., Diamond, B., Farrington, D. P., Tremblay, R. E., Welsh, B. C., & Gonzalez, J. M. R. (2016). A meta-analysis update on the effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology, 12*(2), 229–248. https://doi.org/10.1007/s11292-016-9256-0
Reid, J. B., Patterson, G. R., & Synder, J. (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. American Psychological Association.
Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems – A meta-analytic review. *Journal of Child Psychology and Psychiatry, 47*(1), 99–111.
Reynolds, A. J., Temple, J. A., Ou, S.-R., Arteaga, I. A., & White, B. A. B. (2011). School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. *Science, 333*, 360–364. https://doi.org/10.1126/science.1203618
Rhew, I. C., Hawkins, J. D., Murray, D. M., Fagan, A. A., Oesterle, S., Abbott, R. D., & Catalano, R. F. (2016). Evaluation of community-level effects of community that care on adolescent drug use and delinquency using a repeated cross-sectional design. *Prevention Science, 17*(2), 177–187. https://doi.org/10.1007/s11121-015-0613-4
Romeo, R., Knapp, M., & Scott, S. (2006). Economic cost of severe antisocial behaviour in children—and who pays it. *The British Journal of Psychiatry, 188*(6), 547–553. https://doi.org/10.1192/bjp.bp.104.007625
Rothbaum, F., & Weisz, J. R. (1994). Parental caregiving and child externalizing behavior in nonclinical samples: A meta-analysis. *Psychological Bulletin, 116*(1), 55–74. https://doi.org/10.1037/0033-2909.116.1.55
Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology, 4*(1), 61–81. https://doi.org/10.1007/s11292-007-9046-9
Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The triple P-positive parenting program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychological Review, 34*(4), 337–357. https://doi.org/10.1016/j.cpr.2014.04.003
Sandler, I., Wolchik, S. A., Cruden, G., Mahrer, N. E., Ahn, S., Brincks, A., & Brown, C. H. (2014). Overview of meta-analyses of the prevention of mental health, substance use, and conduct problems. *Annual Review of Clinical Psychology, 10*, 243–273. https://doi.org/10.1146/annurev-clinpsy-050212-185524
Sandler, I. N., Schoenfelder, E. N., Wolchik, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annunal Review of Psychology, 62*, 299–329. https://doi.org/10.1146/annurev.psych.121208.131619
Schoemaker, N. K., Wentholt, W. G. M., Goemans, A., Vermeer, H. J., Juffer, F., & Alink, L. R. A. (2020). A meta-analytic review of parenting interventions in foster care and adoption. *Development and Psychopathology, 32*, 1149–1172. https://doi.org/10.1017/S0054579419000798
Schweinhart, L. J. (2013). Long-term follow-up of a preschool experiment. *Journal of Experimental Criminology, 9*, 389–409. https://doi.org/10.1007/s11292-013-9190-3
Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: Follow-up study of antisocial children into adulthood. *British Medical Journal, 323*, 191. https://doi.org/10.1136/bmj.323.7306.191
Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27*(2), 171–186. https://doi.org/10.1016/S0005-7894(96)80013-X
Sundell, K., Beelmann, A., Hasson, H., & von Thiele Schwarz, U. (2016). Novel programs, international adoptions, or contextual adaptions? Meta-analytical results from german and swedish intervention research. *Journal of Clinical Child and Adolescent Psychology, 45*(6), 784–796. https://doi.org/10.1080/15374416.2015.1020540
Thomas, R., Abell, B., Webb, H. J., Avdagic, E., & Zimmer-Gembeck, M. J. (2017). Parent-child interaction therapy: A meta-analysis. *Pediatrics, 140*(3), 1–2. https://doi.org/10.1542/peds.2017-0352
Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology, 53*(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4
van Aar, J., Leijten, P., de Castro, B. O., & Overbeek, G. (2017). Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behaviour. *Clinical Psychology Review, 51*, 153–163. https://doi.org/10.1016/j.cpr.2016.11.006
van Heel, M., van den Noortgate, W., Bijttebier, P., Colpin, H., Goossens, L., Verschueren, K., & Van Leeuwen, K. (2019). Parenting and externalizing problem behavior in adolescence: Combining the strengths of variable-centered and person-centered approaches. *Developmental Psychology, 55*(3), 653–673. https://doi.org/10.1037/dev0000644
van Ryzin, M. J., Kumpfer, K. L., Fosco, G. M., & Greenberg, M. T. (2016). *Family-based prevention programs for children and adolescents. Theory, research, and large-scale dissemination*. Psychology Press.
Weber, L., Kamp-Becker, I., Christiansen, H., & Mingebach, T. (2019). Treatment of child externalizing behavior problems: A comprehensive review and meta-meta-analysis on effects of parent-based interventions on parental characteristics. *European Child & Adolescent Psychiatry, 28*(8), 1025–1036. https://doi.org/10.1007/s00787-018-1175-3
Weisenmuller, C., & Hilton, D. (2020). Barriers to access, implementation, and utilization of parenting interventions: Considerations for research and clinical applications. *American Psychologist*. https://doi.org/10.1037/amp0000613. Advance online publication.
Weiss, M., Schmucker, M., & Lösel, F. (2015). Meta-analyse zur wirkung familienbezogener präventionsmaßnahmen in deutschland [meta-analytic review of family-based prevention programs in Germany]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 44*(1), 27–44. https://doi.org/10.1026/1616-3443/a000298
Wilson, D. B. (2010). SPSS macros for performing meta-analytic analyses. http://mason.gmu.edu/~dwilsonb/ma.html.
Wilson, P., Rush, R., Hussey, S., Puckering, C., Sim, F., Allely, C. S., Doku, P., McConnachie, A., & Gillberg, C. (2012). How evidence-based is an 'evidence-based parenting program'? A PRISMA systematic review and meta-analysis of triple P. *BMC Medicine, 10*, 130. https://doi.org/10.1186/1741-7015-10130
Wyatt Kaminski, J., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36*(4), 567–589. https://doi.org/10.1007/s10802-007-9201-9