META-ANALYSIS

Check for updates

Individual Participant Data Meta-Analysis: Individual Differences in Mediators of Parenting Program Effects on Disruptive Behavior

Liina Björg Laas Sigurðardóttir, MSc, G.J. Melendez-Torres, DPhil, MPH, RN, Sophia Backhaus, DPhil, Frances Gardner, DPhil, Stephen Scott, PhD, FRCPsych, European Parenting Program Research Consortium, Patty Leijten, PhD

Objective: Although parenting programs are the most widely used approach to reduce disruptive behavior in children, there is a notable lack of understanding of the exact changes in parenting that underlie their effects. Challenges include the frequent use of composite measures of parenting behavior and insufficient power to detect mediation effects and individual differences in these in individual trials.

Method: Individual participant data from 14 European randomized controlled trials of social learning–based parenting programs were pooled to examine which specific parenting behaviors best explain program effects. Participants included 3,252 families with children ages 1 to 13 years. Parental use of praise, tangible rewards, physical discipline, harsh verbal discipline, and not following through on discipline were included as putative mediators. Additionally, the study explored whether subgroups of families showing different mediational pathways exist.

Results: Changes in parenting partially mediated program effects, with all included parenting behaviors except parental use of praise serving as unique mediators. Less harsh verbal discipline and increased following through on discipline were the strongest mediators. The study identified 3 subgroups with distinct responses to parenting programs. Most families benefited, partly through increased following through on discipline; families with the least or most difficulties were more likely to benefit less or not at all.

Conclusion: These findings offer insight into the specific parenting behavior changes key to reducing disruptive child behavior, while highlighting the need for innovative research methodologies to gain a deeper understanding of individual differences in parenting program benefits and mechanisms.

Plain language summary: Parenting programs are the most widely used strategy to reduce children's disruptive behavior. In this analysis of individual data of 3,252 families that participated in 14 European randomized controlled trials, the authors examined which specific changes in parenting behaviors best explain the effects of parenting programs on children's behavior, and whether these changes might differ across different types of families. The authors found that changes in parenting behavior partially explained program effects on disruptive behavior. Reduced harsh verbal discipline and increased following through on discipline seemed the most important parenting behaviors. While most families benefited from parenting programs, those facing the most or the least difficulties benefited less.

Study preregistration information: Study Preregistration: Individual Participant Data Meta-analysis: Individual Differences in Mediators of Parenting Program Effects on Disruptive Behavior; https://doi.org/10.1016/j.jaac.2023.11.005.

Key words: disruptive behavior; individual participant data meta-analysis; mediators; parenting programs

J Am Acad Child Adolesc Psychiatry 2025;64(5):564-576.

Parenting programs are the most widely used strategy for reducing disruptive behavior in children, such as oppositional, aggressive, and rule-breaking behavior.[1,2] However, we still know little about what exact changes in parenting behavior underlie program effects on children's disruptive behavior. In fact, most mediation studies have been unable to identify any parenting behavior as the mechanism of change in parenting programs.[3,4] This is likely, at least in part, because individual trials tend to be underpowered to detect mediation effects[5] and because such trials are unable to take the known heterogeneity in program effectiveness into account.[6] Without an understanding of how the change in children's disruptive behavior comes about, however, important opportunities to improve parenting programs are hampered.[7] By using an unprecedentedly large combined sample from multiple countries and programs, the present study aimed to improve our understanding of how parenting programs for children's disruptive behavior work by identifying their mediators and individual differences between families in these mediators.

Parenting Programs for Children's Disruptive Behavior

Disruptive behavior in childhood predicts poor outcomes in multiple domains of functioning in children[8,9] and is one of the main reasons for referrals to mental health services.[10]

The recommended strategy for prevention and treatment of children's disruptive behavior are parenting programs grounded in social learning theory principles[2] stating that disruptive behavior develops and is maintained by coercive parent–child interaction patterns in which parents and children inadvertently reinforce each other's aversive behavior.[11] For example, a child's defiance with parental requests is reinforced if it results in parents giving in to the child's behavior. Similarly, parental harshness is reinforced if it results in a child's compliance with parental requests. Over time, these reinforcement processes are posited to lead to progressively more dysfunctional parent–child interactions that become increasingly difficult for parents to manage.

Social learning–based parenting programs aim to break coercive interaction patterns by guiding parents to use more effective behavior management strategies, such as reinforcement techniques to increase positive child behavior and nonviolent discipline techniques to reduce disruptive child behavior.[12] By targeting parenting behavior, it is thus believed that parenting programs can lead to use of more effective parenting practices, which, in turn, improve children's behavior. Meta-analyses show that parenting programs have robust small to moderate effects on parenting behavior (eg, $d = -0.18$ to $d = -0.46$ in Backhaus *et al.*[13]) and on children's disruptive behavior (eg, $d = -0.45$ to $d = -0.49$ in Mingebach *et al.*[14]). More recently, the field has moved toward individual participant data meta-analyses (IPDMA). In an IPDMA of the Incredible Years parenting program, Leijten *et al.*[15] pooled data from 1,799 families participating in 14 randomized trials and found that the Incredible Years program reduced children's disruptive behavior and attention-deficity/hyperactivity disorder (ADHD) symptoms as well as parental use of corporal punishment, threats, and shouting. Parental use of praise increased, but children's emotional symptoms and parental use of tangible rewards, monitoring, and laxness did not change.

Mediators of Program Effects

Although effects on both parenting and disruptive child behavior have been established, research has yet to reliably identify which parenting behaviors act as the key mechanisms through which parenting programs work. A review of mediation studies found that the majority of published studies did not find support for parenting behavior as the mediator of parenting program effects on disruptive child behavior.[3] More recent studies also present inconsistent results.[16,17] One reason for this may be that most mediation studies use composite measures of parenting behavior, such as positive parenting and negative parenting. However, exactly what these composite measures comprise typically differs between studies. For example, positive parenting included warmth and involvement in the study by Hanisch *et al.*,[18] but praise and rewards in the study Bjørknes *et al.*[19] A thorough understanding of how parenting programs work requires knowledge of which specific parenting behaviors mediate program effects.

Plausible mediators are commonly targeted parenting behaviors in programs, such as the use of praise and tangible rewards to increase positive child behavior, which may translate into reduced disruptive behavior.[20] Furthermore, reducing physical and harsh verbal discipline is a key goal of most programs because reduced parental harshness can gradually de-escalate parent–child interactions.[12] Finally, not following through on discipline may mediate program effects as programs strive to increase parental consistency to reduce children's noncompliance.[21] Besides being grounded in social learning theory underlying programs,[11] these parenting behaviors have been shown to be affected by such programs in the IPDMA of the Incredible Years parenting program.[15]

In addition to composite measures of parenting, one reason for inconsistent findings may be the substantial statistical power generally required for detecting mediation effects. A review of recent studies in psychology and educational sciences found that studies tend to have insufficient sample sizes to test their mediation models.[22] This discrepancy between the actual and the minimum required sample size for mediation is likely to be even larger in child and adolescent psychiatry, as intervention trials typically include small sample sizes.[5] One way to achieve sufficient statistical power is to pool data from different trials in an IPDMA.[23] Besides increased power and generalizability of findings across programs and populations (eg, in terms of sociodemographic variables, problem severity and complexity), an IPDMA can address conflicting findings identified in a previous review of mediators in parenting programs by estimating mean effects across trials, while minimizing variance due to differences in analytic methods.[24] This is a promising, but still relatively uncommon approach in child and adolescent psychiatry research to examine mediators of program effects.[25]

Individual Differences in Mediators

Another challenge in identifying mediators is that individual studies typically do not consider individual differences between families in program effects. Individual differences in parenting program effects are substantial[6] and may suggest differences in the underlying mediators of program effects. In such cases, conducting mediation analyses for the whole group would likely affect the strength of the mediation effects, possibly even resulting in null effects.[4] It is therefore recommended to combine the study of how psychotherapy programs work with the study of for whom they work.[5,26]

Differences in parenting program effects are rarely explained by individual family characteristics. Traditional moderation analyses have revealed few consistent moderators of parenting program response, with only greater initial severity of child disruptive behavior consistently found to moderate program effects.[27] In contrast to the traditional variable-centered approach to moderation, a person-centered approach can explore how broader family profiles are associated with differential intervention effects. Such mixture analytic models identify subgroups characterized by different intervention response patterns, after which family characteristics can be used to predict membership to these subgroups.[28] For example, a recent trial of the Incredible Years parenting program found that families with high baseline levels of both harsh parenting and child disruptive behavior were most likely to show the strongest decrease in child disruptive behavior.[29] Interestingly, harsh parenting alone was not identified as a moderator in this study when using traditional moderation analyses. Using a more person-centered approach can thus not only take into account the relation between potential moderators, but also reveal moderators not otherwise identified.

Present Study

We examined which specific parenting behaviors mediate parenting program effects on children's disruptive behavior using pooled longitudinal individual participant data. As potential mediators, we included 5 parenting behaviors typically targeted in programs and commonly measured in studies: parental use of praise, tangible rewards, physical discipline, harsh verbal discipline, and not following through on discipline. Due to previous inconsistent findings, we had no a priori hypotheses about which of these parenting behaviors play the biggest role in explaining program effects. Furthermore, we explored individual differences in mediators by testing whether different subgroups of families could be identified for whom parenting programs work through different mediators and which family functioning and sociodemographic characteristics describe these subgroups. As predictors of subgroups, we selected baseline family functioning and sociodemographic variables that have been proposed as predictors or moderators of parenting program effects[30-34] and that were measured in the majority of the included trials. Specifically, we included baseline levels of parental depression; ADHD symptoms and emotional problems in children; age and gender of children; and parental educational level, income, and marital status (single parenthood). We had no a priori hypotheses about which of these variables would predict subgroups with distinct mechanisms because most of these variables have shown different patterns in different samples.

METHOD

Procedure

This preregistered IPDMA (PROSPERO #CRD42022262594) built on a recent systematic review[13] that searched for trials of the effects of parenting programs on disruptive child behavior in 26 online databases (eg, PsycINFO, MEDLINE), 4 trial registries (eg, ClinicalTrials.gov, metaRegister of Controlled Trials), and gray literature. Details of the search terms can be found in Supplement 1, available online. The latest search update occurred in August 2022. We included trials in the present study if they were conducted in Europe; tested in a randomized controlled trial a social learning–based parenting program (>80% of sessions were with parents); included children with a mean age of 2 to 10 years old, as social learning–based programs tend to target this age group; and had at least 2 follow-up assessments, in order to be able to establish the temporal precedence assumed in mediation models. Figure S1, available online, presents the trial selection process. Observational measures were included in only a few trials (14%) and are more difficult to harmonize due to the data being of a different nature, such as frequency counts as opposed to Likert scales. We therefore included only parent reports of parenting and child behavior. Trials were excluded if they specifically targeted special groups, such as children with ADHD or physical, learning, or developmental disabilities or children in temporary foster care, because programs that target these populations tend to include content that addresses their specific needs, which could influence the program's mediators. Based on these criteria, principal investigators of 22 eligible trials were contacted, and 14 (64%) shared their anonymized data. The raw participant-level data were checked for missing, out-of-range, or inconsistent items. Uncertainties were resolved with the corresponding author.

Participants

Data came from 3,252 families (88% mothers), who were diverse in terms of their country of residence (7 European countries), educational level (24% primary or lower secondary education), income (39% low income), marital status (20% single parents), and cultural and ethnic background (eg, in Alsem *et al.*,[35] 76% was born outside of the Netherlands). Families participated in 14 trials on 6 theoretically similar parenting programs (Table 1)[35-48] (Table S1, available online, contains more details). Triple P–Positive Parenting Program was the most commonly evaluated program (k = 6), followed by the Incredible Years (k = 3) and Parent Management Training–Oregon

TABLE 1 Trial Characteristics

Authors, year	N	Country	Intervention	Trial type[a]	Control group	Post-test to FU (mo)	Child age, y		Girls (%)
							Range	(Mean)	
Alsem *et al.*, 2021[35]	75	The Netherlands	IY	Selective	CAU	8	2-10	(6.25)	49
Cina *et al.*, 2011[36]	398	Switzerland	Triple P	Universal	Wait-list	4	2-10	(8.70)	46
Eichelberger *et al.*, 2010[37]	91	Germany	Triple P	Universal	No care	6	3-6	(4.23)	43
Hanisch *et al.*, 2010[38]	122	Germany	PEP	Indicated	No care	3	2-6	(4.58)	28
Kjøbli and Ogden, 2012[39]	190	Norway	Brief Parent Training	Indicated	CAU	6	3-12	(7.22)	32
Kjøbli *et al.*, 2013[40]	120	Norway	PMTO	Treatment	Wait-list	6	2-13	(8.61)	36
Maaskant *et al.*, 2016[41]	86	The Netherlands	PMTO	Indicated	CAU	4	4-11	(7.70)	52
Malti *et al.*, 2011[42]	632	Switzerland	Triple P	Universal	CAU	12	5-8	(7.04)	47
Nogueira *et al.*, 2021[43]	134	Portugal	Triple P	Selective	CAU	6	2-12	(7.13)	55
Sampaio *et al.*, 2015[44]	488	Sweden	Triple P	Universal	Wait-list	6	3-5	(4.11)	47
Scott *et al.*, 2014[45]	214	England	IY	Indicated	CAU	12	3-7	(6.06)	44
Smallegange *et al.*, 2016[46]	80	The Netherlands	Triple P	Selective	CAU	6	1-5	(2.57)	44
van Aar *et al.*, 2021[47]	239	The Netherlands	Positive Notes	Selective	Wait-list	0.5	3-10	(6.19)	37
Weeland *et al.*, 2017[48]	386	The Netherlands	IY	Indicated	CAU	4	3-8	(6.31)	45

Note: *CAU = care-as-usual; FU = follow-up; IY = Incredible Years; PEP = Prevention program for preschool children with Externalizing Problem behavior; PMTO = Parent Management Training–Oregon; Triple P = Triple P–Positive Parenting Program.*

[a]*Trial types: Indicated—families with emerging disruptive child behavior; Selective—families at higher risk for disruptive child behavior (eg, selected based on socioeconomic deprivation/parenting difficulties); Treatment—families referred to outpatient clinics for children's mental health problems; Universal—general community samples.*

(PMTO) (k = 2). Prevention program for preschool children with Externalizing Problem behavior (PEP), Brief Parent Training, and Positive Notes were each evaluated in 1 trial. Trials were conducted in the Netherlands (k = 5); Germany, Norway, and Switzerland (all k = 2); and England, Sweden, and Portugal (all k = 1). Children were between 1 and 13 years of age (mean [SD] age = 6.06 [1.98]), and most were boys (56%). Six trials included only children who scored above a cutoff for disruptive behavior, and 8 were prevention trials. Time between first and second follow-up assessments ranged from 2 weeks to 12 months (mean = 6 months).

Measures

Parenting Behavior (First Follow-up Assessment). We included parental use of praise (9 trials), tangible rewards (10 trials), physical discipline (13 trials), harsh verbal discipline (13 trials), and not following through on discipline (13 trials) as possible mediators. From the parent-reported measures used across trials (Alabama Parenting Questionnaire, Parenting Scale, Parenting Practices Inventory, Parenting Behavior Questionnaire, Nijmegen Parenting Questionnaire, Parental Dimensions Inventory, and a study-specific parenting interview), we selected the items that fit each parenting construct, using established subscales where possible. A similar item-level matching approach has been used successfully in large British cohort studies.[49] All items and their internal consistencies can be found in Supplement 2 and Table S2, available online. We computed an average score for each parenting construct for every participant and converted all response scales to a 1-to-7 scale, the most commonly used scale across trials.[15]

Children's Disruptive Behavior (Second Follow-Up Assessment). The intensity scale of the Eyberg Child Behavior Inventory (ECBI)[50] was used as the primary measure because it was used in 7 of 14 trials. The intensity scale consists of 36 items on the frequency of disruptive behavior of children. Parents report how often each behavior occurs using a 7-point Likert scale (1 = never to 7 = always), and a total sum score is calculated. The scale is widely used to detect change in disruptive behavior in children and has demonstrated good convergent and discriminant validity.[51,52] We used norm deviation scores to convert scores from other measures (aggressive behavior subscale of the Child Behavior Checklist [CBCL], conduct problems subscale of the Strengths and Difficulties Questionnaire [SDQ], and aggressive behavior subscale of the Social Behavior Questionnaire) to ECBI scores.[15]

Predictors of Family Subgroups (Baseline Assessment). Predictors of family subgroup membership included the following:

- Children's emotional problems and ADHD symptoms (8 and 9 trials, respectively): The emotional problems and hyperactivity/inattention subscales of the SDQ were used as the primary measures of children's emotional problems and ADHD symptoms, respectively. Four trials used the internalizing behavior and ADHD subscales of the CBCL. We used norm deviation scores to convert these CBCL scores into SDQ scores.[15]
- Parental depressive symptoms (7 trials): The Depression Anxiety Stress Scales (DASS)[53] was used as the primary measure of parental depressive symptoms. One trial used the General Health Questionnaire,[54] and another used the Brief Symptom Inventory.[55] For these trials, norm deviation scores were used to convert scores into DASS scores.
- Parental educational level (14 trials): Highest level of education completed by the primary caregiver was coded according to International Standard Classification of Education (ISCED-11) categories (1 = primary education or less; 2 = lower secondary education; 3 = upper secondary with or without qualifications; 4 = post-tertiary education; 5 = bachelor-, master-, or doctorate-level education).[56]
- Low income (9 trials): Indicators of low family income, such as having low annual income or receiving financial benefits, were dichotomized (0 = no low income; 1 = low income).
- Single parenthood (14 trials): Primary caregivers who did not live with a partner were coded as single parents (0 = not single parent; 1 = single parent).

Risk of Bias

We assessed the risk of bias within the included trials as high, low, or unclear using the Cochrane risk-of-bias tool for randomized controlled trials[57] (Figure S2, available online).

Statistical Analyses

Each individual participant dataset was prepared for merging in IBM SPSS version 28.0 (IBM Corp., Armonk, NY). Analyses followed a 1-stage IPDMA approach[23] with the clustered nature of the pooled dataset (ie, participants clustered within trials) taken into account by including dummy variables for each trial, as the number of clusters was too low for multilevel modeling or cluster-robust standard errors.[58] All mediation analyses were conducted using path analysis in Mplus version 8.8.[59] First, we tested a mediation model with all 5 parenting behaviors simultaneously included to examine whether parenting program effects on children's disruptive behavior could be explained by changes in parenting behaviors and which parenting behaviors explained the program effects best. Baseline measures of disruptive behavior were included as a covariate in all regression models, with baseline measures of the corresponding parenting behavior additionally included as a covariate in the regression model for each mediator. To evaluate model fit, we examined the root mean square error of approximation[60] and the comparative fit index.[61] A root mean square error of approximation point estimate of <0.05 and a comparative fit index value of >0.95 were used as indicators of good model fit.[62] We calculated bias-corrected bootstrap CIs with 10,000 repetitions to test the indirect effects[63] and examined partially standardized indirect effects to evaluate their magnitude. As there are no guidelines for small, medium, and large standardized indirect effects,[64] we interpreted the magnitude mainly in relation to the other mediators in the model. Although this enabled us to infer which of the parenting behaviors play the biggest role in explaining program effects, it does not allow for conclusions about the absolute magnitude of these effects. Evidence from recent simulation studies suggests the study had enough power to detect even small mediation effects.[22] Missing values were handled using , so that all available information from all trials was used to estimate mediators across trials.

In a second step, a mixture mediation model was used to identify latent subgroups of families that show different mediational pathways. All regression paths in the previous mediation model except covariate paths were allowed to vary across the latent subgroups. The Bayesian information criterion[65] and the bootstrapped likelihood ratio test[66] were evaluated to identify the optimal number of latent subgroups, as recommended for mixture mediation models.[67] Additionally, we considered subgroup proportions, classification efficiency, and subgroup interpretability. Furthermore, baseline family functioning and sociodemographic variables were used to predict latent subgroup membership by conducting multinomial logistic regression analyses using a 3-step classification (R3STEP) procedure in Mplus.[68] As R3STEP uses listwise deletion to handle missing data in auxiliary variables, we inserted each predictor individually in the model.

RESULTS

Preliminary Results

Means and standard deviations for study variables across time points are presented in Table 2. Children's disruptive behavior at baseline ranged from minimal or no symptoms (ECBI score of 38) to severe clinical levels (ECBI score of 245; highest possible ECBI score is 252), reflecting the

TABLE 2 Descriptive Statistics for Study Variables

			Parenting intervention						Control					
			Baseline		First follow-up		Second follow-up		Baseline		First follow-up		Second follow-up	
Variable	**k**	**n**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Outcome variable														
Disruptive behavior	14	3,052	118.19	(30.90)	109.44	(26.50)	108.17	(26.16)	116.29	(30.77)	111.06	(27.72)	109.21	(26.54)
Mediators														
Praise	9	2,433	5.62	(1.07)	5.71	(1.04)	—		5.62	(1.15)	5.54	(1.15)	—	
Tangible rewards	10	2,519	3.46	(1.64)	3.64	(1.57)	—		3.47	(1.66)	3.40	(1.55)	—	
Physical discipline	13	2,819	1.96	(1.05)	1.77	(0.96)	—		1.89	(1.04)	1.75	(0.96)	—	
Harsh verbal discipline	13	3,102	3.08	(1.27)	2.70	(1.15)	—		3.10	(1.28)	2.99	(1.20)	—	
Not following through	13	3,083	2.99	(1.11)	2.67	(1.05)	—		2.97	(1.15)	2.84	(1.09)	—	
Predictors of subgroups														
ADHD symptoms	9	2,097	3.91	(2.77)	—		—		3.88	(2.76)	—		—	
Emotional problems	11	2,551	1.94	(1.98)	—		—		1.88	(1.93)	—		—	
Parental depressive symptoms	7	1,454	5.72	(7.01)	—		—		5.91	(7.19	—		—	

Note: *The number of participants reported for each variable represent the number of participants at baseline. ADHD = attention-deficit/hyperactivity disorder.*

diverse range of families targeted in the included trials. Nearly one-fifth (18%) of the children scored above the 90th percentile on disruptive behavior.[69] Baseline scores for harsh verbal discipline, not following through on discipline, and tangible rewards were relatively normally distributed across the possible range. Although baseline scores for physical discipline and praise were slightly positively and negatively skewed, respectively, both distributions encompassed the entire range of possible scores, and skewness was below the absolute value of 2.

Mediators of Parenting Program Effects

The path model yielded poor fit to the data, as expected when including multiple dummy variables to account for clustering of participants within trials. Sensitivity analyses with cluster-robust standard errors yielded good model fit and similar results, and inspection of modification indices further indicated that the model misfit was indeed due to the dummy variables (see Tables S3 and S4, available online, for model fit and modification indices). Because accounting for within-trial clustering is essential in an IPDMA,[23] the dummy variables were kept in the model without further model modifications.

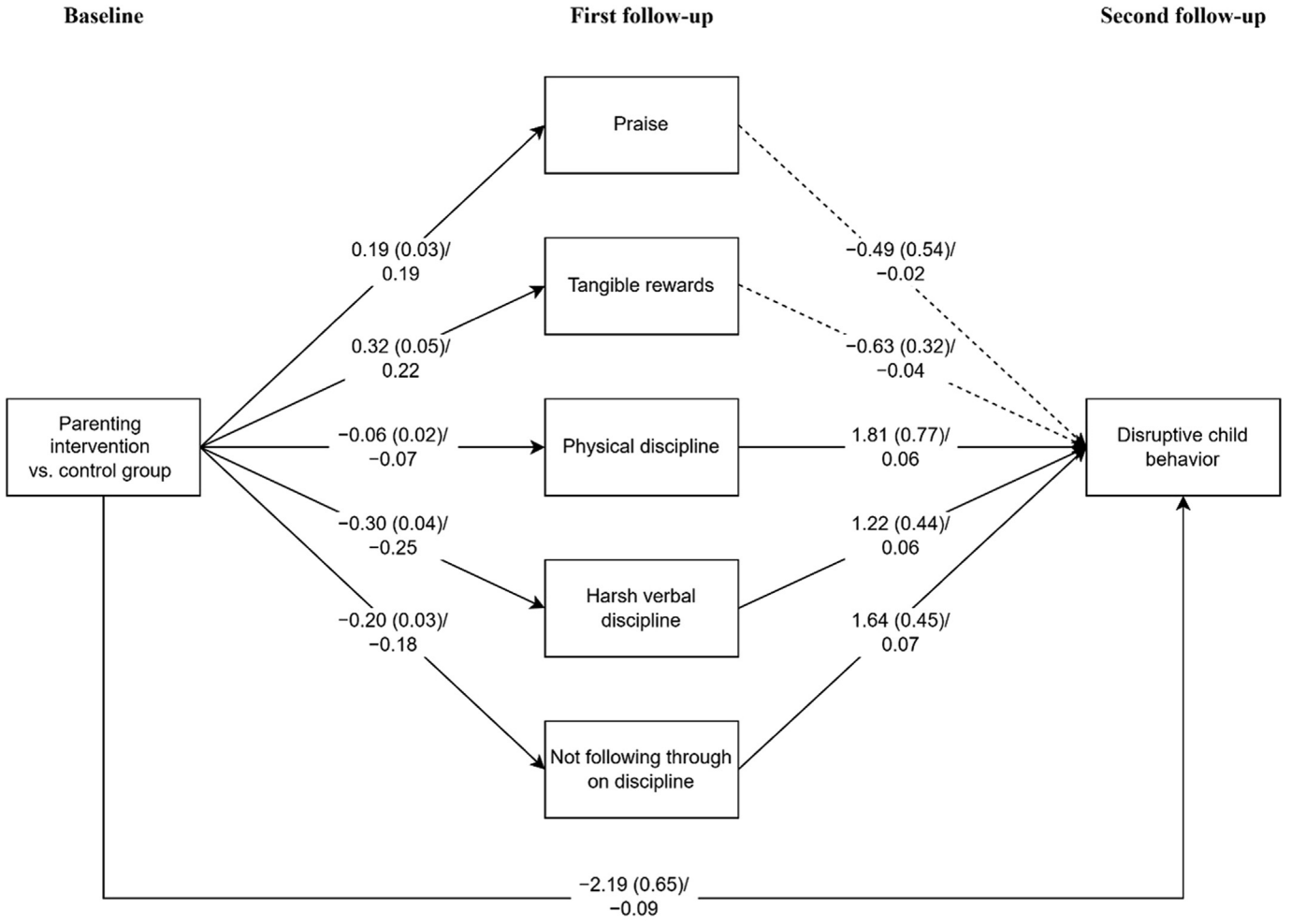
There was a significant total indirect effect of parenting programs on child disruptive behavior through parenting behaviors ($\beta_{\text{indirect}} = -.04$, 95% CI [−0.06, −0.03]) as well as a significant direct effect ($\beta_{\text{direct}} = -.09$, 95% CI [−0.14, −0.04]). In other words, program effects on child disruptive behavior were partly explained by changes in parenting. In terms of individual parenting behaviors, changes in use of tangible rewards ($\beta_{\text{indirect}} = -.008$, 95% CI [−0.017, −0.001]), physical discipline ($\beta_{\text{indirect}} = -.004$, 95% CI [−0.011, −0.001]), harsh verbal discipline ($\beta_{\text{indirect}} = -.014$, 95% CI [−0.027, −0.004]), and not following through on discipline ($\beta_{\text{indirect}} = -.012$, 95% CI [−0.021, −0.005]) uniquely mediated program effects on child disruptive behavior. Changes in parental praise did not uniquely mediate program effects.

Direct effects are shown in Figure 1. Parents randomly assigned to a parenting program reported less use of physical discipline, harsh verbal discipline, and not following through on discipline at the first follow-up assessment, which, in turn, predicted less child disruptive behavior at the second follow-up. Parents also reported increased use of tangible rewards and praise at the first follow-up, but these parenting behaviors did not predict later child disruptive behavior, despite use of tangible rewards also mediating program effects.

Individual Differences in Mediators

To identify the best-fitting model for individual differences in mediators, we tested models with 1 to 5 classes. As both

FIGURE 1 Results From the Overall Mediation Model



Note: *Unstandardized coefficients and standard errors are presented above standardized coefficients in the figure. Significant paths are represented by solid lines, and nonsignificant paths are represented by dashed lines. Baseline measures of parenting behavior and disruptive child behavior are not depicted but were included as covariates in the model.*

Bayesian information criterion and bootstrapped likelihood ratio test values suggested small, incremental improvements in model fit with each added class (Table S5, available online), we selected the most parsimonious model with meaningfully distinct profiles. Whereas the 2- and 3-class models each introduced a new conceptually distinct profile, the 4- and 5-class models included small additional profiles not meaningfully different from the other profiles. Consequently, we selected the 3-class model, with an entropy of 0.84, indicating a good chance of correctly classifying a family to a subgroup.

The first and largest subgroup (69% of families, n = 2,239) consisted of families with elevated levels of baseline child disruptive behavior (ECBI mean score = 121) who responded well to the parenting programs. Descriptive statistics for each subgroup are shown in Table 3. The intervention reduced disruptive child behavior partly through changes in parenting ($\beta_{\text{indirect}} = -.05$, 95% CI [−0.08, −0.03], $\beta_{\text{direct}} = -.10$, 95% CI [−0.16, −0.04]). Of the individual parenting behaviors, only not following through on discipline uniquely mediated program effects ($\beta_{\text{indirect}} = -.02$, 95% CI [−0.03, −0.01]). Within this subgroup, each of the parenting behaviors changed in the expected direction. Table S6 and Figures S3 to S5, available online, present full model results for each subgroup.

The second subgroup (28% of families, n = 897) consisted of families who showed the lowest levels of baseline disruptive behavior (ECBI mean score = 105) and did not seem to benefit much from parenting programs. There was no effect on children's disruptive behavior, neither direct nor indirect through parenting. While parents in the intervention group followed through more on

TABLE 3 Descriptive Statistics for Family Subgroups

	Family subgroup					
	Subgroup 1 (n = 2,239 [69%])		Subgroup 2 (n = 897 [28%])		Subgroup 3 (n = 106 [3%])	
Baseline measure	**Mean**	**(SD)**	**Mean**	**(SD)**	**Mean**	**(SD)**
Disruptive child behavior	120.78	(38.71)	105.24	(38.85)	138.77	(43.00)
Praise	5.47	(1.70)	5.91	(1.11)	5.68	(3.14)
Tangible rewards	3.59	(3.45)	3.00	(5.21)	4.04	(3.07)
Physical discipline	1.42	(1.14)	2.81	(2.94)	3.98	(3.09)
Harsh verbal discipline	3.14	(1.47)	2.90	(1.68)	3.98	(2.54)
Not following through on discipline	2.96	(1.70)	3.03	(1.71)	3.24	(1.42)

discipline at the first follow-up assessment, participation in the parenting program did not influence any of the other parenting behaviors.

A small third subgroup (3% of families, n = 106) consisted of higher-risk families who did not respond to parenting programs. There was no effect on either child disruptive behavior or any of the parenting behaviors. These families showed the highest baseline levels of child disruptive behavior (ECBI mean score = 139) as well as the highest levels of physical and harsh verbal discipline.

Although most families belonged to the first subgroup, some families were more likely than others to belong to the second or third subgroup. In fact, 7 of the 8 child and family characteristics tested predicted subgroup membership (all but child gender) (Figure 2; Table S7, available online). Families more likely to belong to subgroup 2, relative to subgroup 1, were those facing fewer stressors at the child and family level. Specifically, this likelihood was larger for families with higher educational levels (odds ratio [OR] = 1.12, 95% CI [1.04, 1.21]) and smaller for families with single parents (OR = 0.52, 95% CI [0.40, 0.67]) and lower incomes (OR = 0.32, 95% CI [0.25, 0.40]) as well as families with children who had more emotional problems (OR = 0.75, 95% CI [0.69, 0.81]). Younger child age was also associated with a smaller likelihood of belonging to subgroup 2 (OR = 0.60, 95% CI [0.56, 0.65]).

In contrast, families more likely than others to belong to subgroup 3, relative to subgroup 1, were those facing more stressors at the child and parent level and some stressors, but not others, at the broader family level. Specifically, likelihood of belonging to subgroup 3 was larger for families with children who exhibited more ADHD symptoms (OR = 1.21, 95% CI [1.13, 1.29]) or emotional problems (OR = 1.11, 95% CI [1.03, 1.20]) and parents who experienced more depressive symptoms (OR = 1.07, 95% CI [1.04, 1.10]). Furthermore, the likelihood was smaller for families with higher educational levels (OR = 0.62, 95% CI [0.53, 0.72]) and single parents (OR = 0.48, 95% CI [0.25, 0.94]). Similar to subgroup 2, younger child age (OR = 0.57, 95% CI [0.46, 0.71]) was also associated with a smaller likelihood of belonging to subgroup 3.

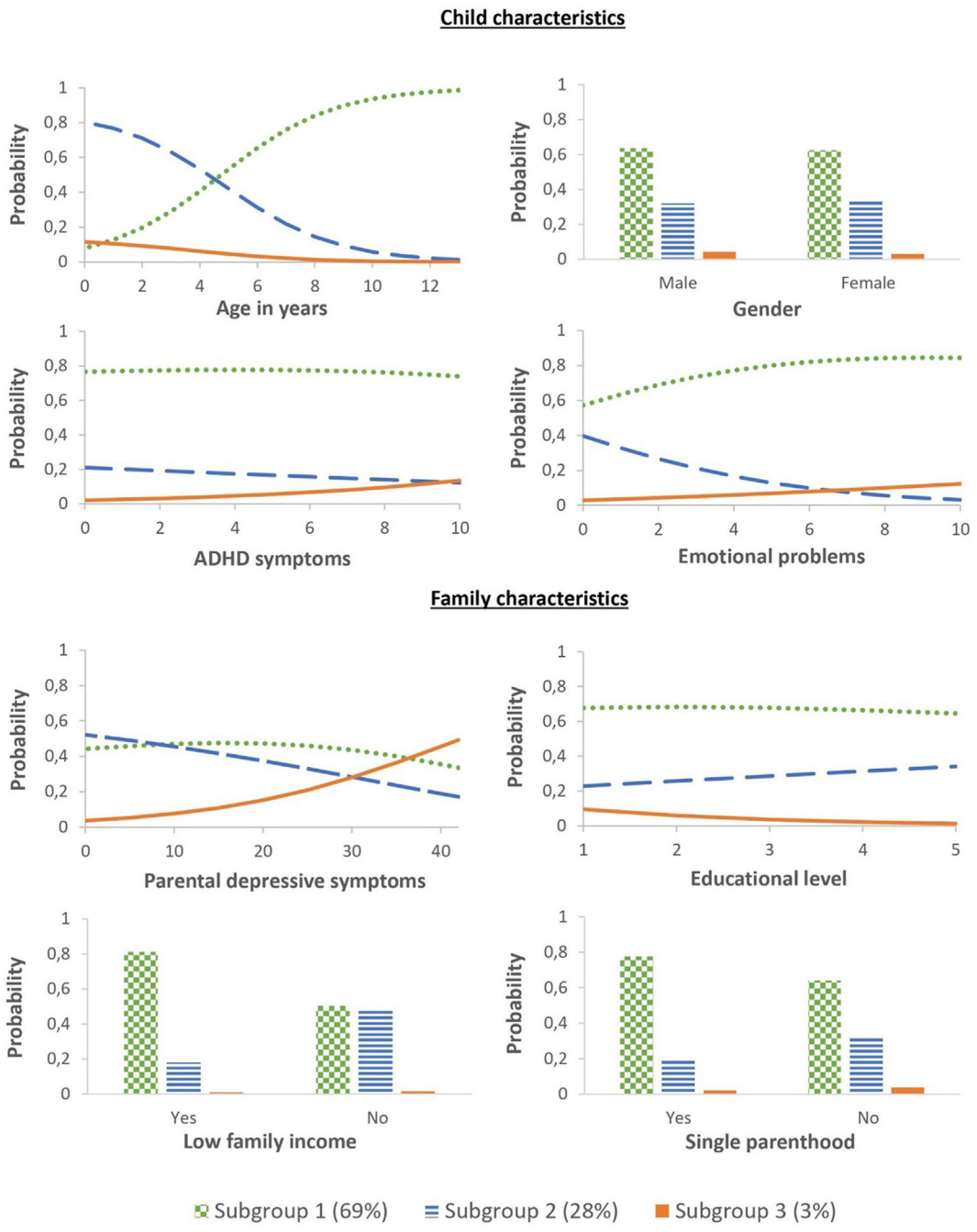
DISCUSSION

We found that of the 5 putative mediators tested, the following 4 uniquely mediated parenting program effects on disruptive child behavior: increased use of tangible rewards and less physical discipline, harsh verbal discipline, and not following through on discipline. We identified 3 subgroups of families, each with their own response to parenting programs. The largest subgroup showed decreases in child disruptive behavior, explained partly by parents following through more on discipline. Two other subgroups both showed no effects on disruptive behavior and minimal or no effects on parenting behavior, respectively.

Our first finding suggests that reductions in harsh verbal and physical discipline and increases in following through on discipline and the use of tangible rewards each play a unique role in reducing disruptive child behavior. This is in line with the core content of the parenting programs, which guides parents to shift their attention to reward positive child behavior and use nonviolent disciplining techniques to deal with disruptive child behavior.[12] Parental use of praise did not uniquely mediate program effects. Although there is evidence that increased praise can reduce disruptive child behavior,[70] our results suggest that changes in praise are not essential above and beyond the other changes in parenting behavior.

Reduced harsh verbal discipline and increased following through on discipline were the strongest mediators. This aligns with the theory of change underlying behavioral

FIGURE 2 Probabilities of Subgroup Membership Conditional on Child and Family Characteristics



parenting programs that breaking coercive parent–child interaction cycles is key to reducing disruptive child behavior.[11] Importantly, however, the included parenting behaviors only partially mediated program effects. The relatively strong remaining direct effect suggests that there are likely additional changes in family dynamics contributing to program effects on disruptive behavior. One main candidate may be parental warmth and responsivity, which increases positivity and security in the parent–child relationship and fosters the development of the child's emotion and behavior regulation.[71] Many established parenting programs, including the programs in this study, aim to improve the parent–child relationship before introducing behavior management techniques.[12] Only 4 of the trials measured parental warmth, however—too few to reliably examine this mediator in our study. To better understand

the mechanisms of change of programs, more studies are needed that include warmth, as well as other theoretically compelling, but less commonly measured potential mediators.

Whereas earlier studies have consistently suggested that families with more severe baseline disruptive child behavior benefit more from parenting programs,[27,72] our subgroup findings revealed a nonlinear relation between program benefits and baseline severity and risk. Both families with relatively good baseline functioning and families with more severe and complex (ie, comorbid) child mental health problems and broader family difficulties (eg, parental depression and less education) showed little to no benefits. This discrepancy may be explained by our person-centered approach of simultaneously examining a range of family functioning and sociodemographic characteristics. This approach tries to do justice to the fact that risk factors often co-occur and may interact in impacting intervention response. Furthermore, our large and diverse sample may have enabled the identification of this small high-risk subgroup not found in studies employing a person-centered approach in a single sample.[28,29]

Our results have several implications for clinical practice. First, practitioners can perhaps monitor changes in parental harsh verbal discipline and following through on discipline as indicators of meaningful program effects. To be sure, our findings do not necessarily imply that sessions targeting these behaviors are the most essential, as such changes could also be driven by other sessions. Instead, our findings offer insight into the specific changes in parenting that can indicate progress. Furthermore, in line with earlier studies,[28,29] social learning–based parenting programs seem to benefit families most who show risk in areas targeted by these programs—parenting and child behavior problems. This suggests that a carefully targeted approach may be most beneficial. Finally, our results indicate that for some families who face a cumulation of risk factors, the likelihood of benefiting from social learning–based parenting programs might be smaller, and more specialized services may be needed.

Our IPDMA used an unprecedentedly large sample with 3 measurement waves to examine mediators of parenting program effects and represents a novel approach to study individual differences in treatment mechanisms in the field. Similar data pooling approaches could potentially be used to answer other questions in child and adolescent psychiatry research that cannot be addressed in single studies or smaller samples. Moreover, rather than using broad, composite measures of parenting, we tested more targeted measures of parenting behavior as possible mediators of program effects and examined these mediators simultaneously in 1 model to disentangle the unique mediating effects of each parenting behavior. The diverse study sample, including 6 social learning–based parenting programs in 7 countries, further increases the generalizability of our findings. That said, our exclusion of trials targeting specific groups, such as children with ADHD or developmental disabilities, means we cannot generalize our findings to these groups.

Some additional limitations also warrant mention. First, data harmonization requires making assumptions, in part because there are no validated subscales for the discrete parenting behaviors (eg, praise and incentives are typically combined into 1 subscale). We therefore used item-level matching to select items for inclusion in the parenting behavior constructs for each trial. Although similar approaches have successfully been used in other studies,[15,49] and item and scale score analyses provide preliminary evidence for adequate reliability and validity, we rely on the assumption that the different measures indeed capture the same construct. Second, our analyses were limited by the pooled data available. The study relied on parent reports for both parenting and child behavior, but such reports are susceptible to various biases, including social desirability bias, and may result in inflated associations between variables due to shared method variance. Furthermore, our study included only pre- and post-intervention assessments because no assessments during the intervention were available in the individual trials. However, as most changes in parenting and child behavior occur during the intervention, such repeated measures might provide important further insight into the actual processes by which these interventions work and shed light on any transactional parent–child processes.[73]

A deeper understanding of how parenting programs work for individual families is essential to better tailor and target programs. Our findings call for the use of innovative research methodologies to further examine individual differences in parenting program benefits and mechanisms, including innovation in the aspects of parenting that are measured and the timing of these assessments. Such efforts have the potential to optimize program benefits for families across the continuum of risk.

CRediT authorship contribution statement

Liina Björg Laas Sigurðardóttir: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **G.J. Melendez-Torres:** Writing – review & editing, Validation, Methodology, Data curation. **Sophia Backhaus:** Writing – review & editing, Resources, Investigation. **Frances Gardner:** Writing – review & editing, Supervision. **Stephen Scott:** Writing – review & editing, Investigation. **European Parenting Program Research Consortium:** Writing – review & editing, Investigation, Resources. **Patty Leijten:** Writing – original draft,

Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Accepted October 3, 2024.

Liina Björg Laas Sigurðardóttir, Sophia Backhaus, and Patty Leijten are with the University of Amsterdam, Amsterdam, the Netherlands. G.J. Melendez-Torres is with the University of Exeter, Exter, United Kingdom. Frances Gardner is with the University of Oxford, Oxford, United Kingdom. Stephen Scott is with King's College London, London, United Kingdom.

The study is supported by a Dutch Research Council Vidi grant (NWO #VI.V-idi.201.065) awarded to P.L. and a grant from the National Institute for Health Research (NIHR) Public Health Research (grant 12-3070-04) awarded to F.G. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the UK Department of Health. The funding source had no role in the design and conduct of the study; preparation, review, or approval of the manuscript; or the decision to submit it for publication.

The research was performed with permission from the Internal Review Board of the Faculty of Social and Behavioural Sciences of the University of Amsterdam.

This work has been prospectively registered: https://doi.org/10.1016/j.jaac.2023.11.005.

G.J. Melendez-Torres served as the statistical expert for this research.

The authors would like to thank the members of the European Parenting Program Research Consortium for contributing their data to the project.

Members of the consortium are, in alphabetical order: Guy Bodenmann, PhD, of the University of Zurich; Ana Catarina Canário, PhD, of the University of Porto; Manfred Döpfner, PhD, of the University of Cologne; Manuel Eisner, PhD, of the University of Cambridge and the University of Zurich; Suzanne de Jong, PhD, of the VU Amsterdam; John Kjøbli, PhD, of the University of Oslo; Ankie Menting, PhD, of Utrecht University; Geertjan Overbeek, PhD, of the University of Amsterdam; Julia Plück, PhD, of the University of Cologne; Raziye Salari, PhD, of Uppsala University; Eva Smallegange, PhD, of War Child Holland; Jolien van Aar, PhD, of VanMontfoort; and Floor van Rooij, PhD, of the University of Amsterdam. In addition, the authors gratefully acknowledge the researchers, practitioners, and families who participated in each of the original trials, and the funders of the original trials.

Disclosure: Julia Plück, of the European Parenting Program Research Consortium, has disclosed interest in relation to co-authorship of the parenting program "the Prevention program for preschool children with Externalizing Problems (PEP)." The other members of the consortium have reported no biomedical financial interests or potential conflicts of interest. Liina Björg Laas Sigurðardóttir, G.J. Melendez-Torres, Sophia Backhaus, Frances Gardner, Stephen Scott, European Parenting Program Research Consortium, and Patty Leijten have reported no biomedical financial interests or potential conflicts of interest.

Correspondence to Liina Björg Laas Sigurðardóttir, MSc, University of Amsterdam, Postbus 15776, 1001 NG Amsterdam, the Netherlands; e-mail: l.bjorglaassigurdardottir@uva.nl

0890-8567/$36.00/©2024 American Academy of Child and Adolescent Psychiatry. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.1016/j.jaac.2024.10.003

REFERENCES

1. Beelmann A, Arnold LS, Hercher J. Parent training programs for preventing and treating antisocial behavior in children and adolescents: a comprehensive meta-analysis of international studies. Aggress Violent Behav. 2022;68:101798. https://doi.org/10.1016/j.avb.2022.101798
2. Weisz JR, Kazdin AE, eds. Evidence-Based Psychotherapies for Children and Adolescents. 3rd ed. New York, NY: Guilford Press; 2017.
3. Forehand R, Lafko N, Parent J, Burt KB. Is parenting the mediator of change in behavioral parent training for externalizing problems of youth? Clin Psychol Rev. 2014;34(8):608-619. https://doi.org/10.1016/j.cpr.2014.10.001
4. Kazdin AE. Mediators and mechanisms of change in psychotherapy research. Annu Rev Clin Psychol. 2007;3(1):1-27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432
5. Maric M, Wiers RW, Prins PJM. Ten ways to improve the use of statistical mediation analysis in the practice of child and adolescent treatment research. Clin Child Fam Psychol Rev. 2012;15(3):177-191. https://doi.org/10.1007/s10567-012-0114-y
6. Shelleby EC, Shaw DS. Outcomes of parenting interventions for child conduct problems: A review of differential effectiveness. Child Psychiatry Hum Dev. 2014;45(5):628-645. https://doi.org/10.1007/s10578-013-0431-5
7. Holmes EA, Ghaderi A, Harmer CJ, *et al.* The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. Lancet Psychiatry. 2018;5(3):237-286. https://doi.org/10.1016/s2215-0366(17)30513-8
8. Campbell SB, Spieker S, Burchinal M, Poe MD. The NICHD Early Child Care Research Network. Trajectories of aggression from toddlerhood to age 9 predict academic and social functioning through age 12: Outcomes of aggression trajectories. J Child Psychol Psychiatry. 2006;47(8):791-800. https://doi.org/10.1111/j.1469-7610.2006.01636.x
9. Loeber R, Burke JD, Pardini DA. Development and etiology of disruptive and delinquent behavior. Annu Rev Clin Psychol. 2009;5(1):291-310. https://doi.org/10.1146/annurev.clinpsy.032408.153631
10. Kimonis ER, Frick PJ, McMahon RJ. Conduct and oppositional defiant disorders. In: Mash EJ, Barkley RA, eds. Child Psychopathology. New York, NY: Guilford Press; 2014:145-179.
11. Patterson GR. Coercive Family Processes. Zug, Switzerland: Castalia; 1982.
12. Kaehler LA, Jacobs M, Jones DJ. Distilling common history and practice elements to inform dissemination: Hanf-model BPT programs as an example. Clin Child Fam Psychol Rev. 2016;19(3):236-258. https://doi.org/10.1007/s10567-016-0210-5
13. Backhaus S, Leijten P, Jochim J, Melendez-Torres GJ, Gardner F. Effects over time of parenting interventions to reduce physical and emotional violence against children: a systematic review and meta-analysis. EClinicalMedicine. 2023;60:102003. https://doi.org/10.1016/j.eclinm.2023.102003
14. Mingebach T, Kamp-Becker I, Christiansen H, Weber L. Meta-meta-analysis on the effectiveness of parent-based interventions for the treatment of child externalizing behavior problems. PLoS One. 2018;13(9):e0202855. https://doi.org/10.1371/journal.pone.0202855
15. Leijten P, Gardner F, Landau S, *et al.* Research Review: Harnessing the power of individual participant data in a meta-analysis of the benefits and harms of the Incredible Years parenting program. J Child Psychol Psychiatry. 2018;59(2):99-109. https://doi.org/10.1111/jcpp.12781
16. Fagan AA, Benedini KM. How do family-focused prevention programs work? A review of mediating mechanisms associated with reductions in youth antisocial behaviors. Clin Child Fam Psychol Rev. 2016;19(4):285-309. https://doi.org/10.1007/s10567-016-0207-0
17. Weeland J, Chhangur RR, Jaffee SR, *et al.* Does the Incredible Years reduce child externalizing problems through improved parenting? The role of child negative affectivity and serotonin transporter linked polymorphic region (5-HTTLPR) genotype. Dev Psychopathol. 2018;30(1):93-112. https://doi.org/10.1017/s0954579417000499
18. Hanisch C, Hautmann C, Plück J, Eichelberger I, Döpfner M. The prevention program for externalizing problem behavior (PEP) improves child behavior by reducing negative parenting: analysis of mediating processes in a randomized controlled trial. J Child Psychol Psychiatry. 2014;55(5):473-484. https://doi.org/10.1111/jcpp.12177
19. Bjørknes R, Kjøbli J, Manger T, Jakobsen R. Parent training among ethnic minorities: parenting practices as mediators of change in child conduct problems. Fam Relat. 2012;61(1):101-114. https://doi.org/10.1111/j.1741–3729.2011.00683.x
20. Leijten P, Gardner F, Melendez-Torres GJ, *et al.* Meta-analyses: key parenting program components for disruptive child behavior. J Am Acad Child Adolesc Psychiatry. 2019;58(2):180-190. https://doi.org/10.1016/j.jaac.2018.07.900
21. Kaminski JW, Valle LA, Filene JH, Boyle CL. A meta-analytic review of components associated with parent training program effectiveness. J Abnorm Child Psychol. 2008;36(4):567-589. https://doi.org/10.1007/s10802-007-9201-9
22. Sim M, Kim SY, Suh Y. Sample size requirements for simple and complex mediation models. Educ Psychol Meas. 2022;82(1):76-106. https://doi.org/10.1177/00131644211003261

23. Debray TPA, Moons KGM, Valkenhoef G, *et al*. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. Res Synth Methods. 2015;6(4): 293-309. https://doi.org/10.1002/jrsm.1160
24. Ng MY, DiVasto KA, Cootner S, Gonzalez NAR, Weisz JR. What do 30 years of randomized trials tell us about how psychotherapy improves youth depression? A systematic review of candidate mediators. Clin Psychol (New York). 2020:e12367. https://doi.org/10.1111/cpsp.12367
25. Weisz JR, Venturo-Conerly KE, Fitzpatrick OM, Frederick JA, Ng MY. What four decades of meta-analysis have taught us about youth psychotherapy and the science of research synthesis. Annu Rev Clin Psychol. 2023;19(1):79-105. https://doi.org/10.1146/annurev-clinpsy-080921-082920
26. Webb CA, Hirshberg MJ, Gonzalez O, Davidson RJ, Goldberg SB. Revealing subgroup-specific mechanisms of change via moderated mediation: a meditation intervention example. J Consult Clin Psychol. 2024;92(1):44-53. https://doi.org/10.1037/ccp0000842
27. McMahon RJ, Goulter N, Frick PJ. Moderators of psychosocial intervention response for children and adolescents with conduct problems. J Clin Child Adolesc Psychol. 2021; 50(4):525-533. https://doi.org/10.1080/15374416.2021.1894566
28. Piehler TF, Zhang J, Bloomquist ML, August GJ. Parent and child risk profiles as predictors of response to a conduct problem preventive intervention. Prev Sci. 2022; 23(7):1308-1320. https://doi.org/10.1007/s11121-022-01374-4
29. van Aar J, Leijten P, Orobio de Castro B, *et al*. Families who benefit and families who do not: integrating person- and variable-centered analyses of parenting intervention responses. J Am Acad Child Adolesc Psychiatry. 2019;58(10):993-1003. https://doi.org/10.1016/j.jaac.2019.02.004
30. Florean IS, Dobrean A, Păsărelu CR, Georgescu RD, Milea I. The efficacy of internet-based parenting programs for children and adolescents with behavior problems: a meta-analysis of randomized clinical trials. Clin Child Fam Psychol Rev. 2020;23(4):510-528. https://doi.org/10.1007/s10567-020-00326-0
31. Gardner F, Leijten P, Harris V, *et al*. Equity effects of parenting interventions for child conduct problems: a pan-European individual participant data meta-analysis. Lancet Psychiatry. 2019;6(6):518-527. https://doi.org/10.1016/s2215-0366(19)30162-2
32. Groenman AP, Hornstra R, Hoekstra PJ, *et al*. An individual participant data meta-analysis: behavioral treatments for children and adolescents with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry. 2022;61(2):144-158. https://doi.org/10.1016/j.jaac.2021.02.024
33. Leijten P, Gardner F, Melendez-Torres GJ, *et al*. Co-occurring change in children's conduct problems and maternal depression: latent class individual participant data meta-analysis of the Incredible Years parenting program. Dev Psychopathol. 2019;31(5):1851-1862. https://doi.org/10.1017/s0954579419001068
34. Lundahl B, Risser H, Lovejoy M. A meta-analysis of parent training: moderators and follow-up effects. Clin Psychol Rev. 2006;26(1):86-104. https://doi.org/10.1016/j.cpr.2005.07.004
35. Alsem SC, Menting ATA, De Castro BO, Matthys W. Follow-up effects in a parent-training trial for mothers being released from incarceration and their children. Psychol Public Policy Law. 2021;27(3):421-431. https://doi.org/10.1037/law0000305
36. Cina A, Röösli M, Schmid H, *et al*. Enhancing positive development of children: effects of a multilevel randomized controlled intervention on parenting and child problem behavior. Fam Sci. 2011;2(1):43-57. https://doi.org/10.1080/19424620.2011.601903
37. Eichelberger I, Plück J, Hanisch C, Hautmann C, Jänen N, Döpfner M. Effekte universeller Prävention mit dem Gruppenformat des Elterntrainings Triple P auf das kindliche Problemverhalten, das elterliche Erziehungsverhalten und die psychische Belastung der Eltern. Z Klin Psychol Psychother. 2010;39(1):24-32. https://doi.org/10.1026/1616-3443/a000006
38. Hanisch C, Freund-Braier I, Hautmann C, *et al*. Detecting effects of the indicated prevention programme for externalizing problem behaviour (PEP) on child symptoms, parenting, and parental quality of life in a randomized controlled trial. Behav Cogn Psychother. 2010;38(1):95-112. https://doi.org/10.1017/S1352465809990440
39. Kjøbli J, Ogden T. A randomized effectiveness trial of brief parent training in primary care settings. Prev Sci. 2012;13:616-626. https://doi.org/10.1007/s11121-012-0289-y
40. Kjøbli J, Hukkelberg S, Ogden T. A randomized trial of group parent training: reducing child conduct problems in real-world settings. Behav Res Ther. 2013;51(3):113-121. https://doi.org/10.1016/j.brat.2012.11.006
41. Maaskant AM, van Rooij FB, Overbeek GJ, Oort FJ, Hermanns JM. Parent training in foster families with children with behavior problems: follow-up results from a randomized controlled trial. Child Youth Serv Rev. 2016;70:84-94. https://doi.org/10.1016/j.childyouth.2016.09.005
42. Malti T, Ribeaud D, Eisner MP. The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: a cluster randomized controlled trial. J Clin Child Adolesc Psychol. 2011;40(5):677-692. https://doi.org/10.1080/15374416.2011.597084
43. Nogueira S, Abreu-Lima I, Canário C, Cruz O. Group Triple P—a randomized controlled trial with low-income mothers. Child Youth Serv Rev. 2021;121:105862. https://doi.org/10.1016/j.childyouth.2020.105862
44. Sampaio F, Sarkadi A, Salari R, Zethraeus N, Feldman I. Cost and effects of a universal parenting programme delivered to parents of preschoolers. Eur J Public Health. 2015; 25(6):1035-1042. https://doi.org/10.1093/eurpub/ckv106
45. Scott S, Sylva K, Beckett C, *et al*. Which type of parenting programme best improves child behaviour and reading? Follow-up of the helping children achieve trial. London, UK: Nuffield Foundation; 2014.
46. Smallegange ES, Hermanns JMA, Oort FJ. Evaluating the effectiveness of combining Home-Start and Triple P parenting support in the Netherlands. Child Youth Serv Rev. 2016;68:178-186. https://doi.org/10.1016/j.childyouth.2016.07.009
47. van Aar J, Leijten P, Overbeek G, Thomaes S, Rothman AJ. Does setting goals enhance parenting intervention effects? A field experiment. Behav Ther. 2021;52(2):418-429. https://doi.org/10.1016/j.beth.2020.05.004
48. Weeland J, Chhangur RR, van der Giessen D, Matthys W, de Castro BO, Overbeek G. Intervention effectiveness of The Incredible Years: New insights into sociodemographic and intervention-based moderators. Behav Ther. 2017;48(1):1-18. https://doi.org/10.1016/j.beth.2016.08.002
49. McElroy E, Villadsen A, Patalay P, *et al*. Harmonisation and Measurement Properties of Mental Health Measures in Six British Cohorts. London, UK: CLOSER; 2020.
50. Eyberg SM, Pincus D. Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory–Revised: Professional manual. Odessa, FL: Psychological Assessment Resources; 1999.
51. Boggs SR, Eyberg S, Reynolds LA. Concurrent validity of the Eyberg child behavior inventory. J Clin Child Psychol. 1990;19(1):75-78. https://doi.org/10.1207/s15374424jccp1901_9
52. Rich BA, Eyberg SM. Accuracy of assessment: the discriminative and predictive power of the Eyberg Child Behavior Inventory. Ambul Child Health. 2001;7(3-4):249-257. https://doi.org/10.1046/j.1467-0658.2001.00141.x
53. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther. 1995;33(3):335-343. https://doi.org/10.1016/0005-7967(94)00075-u
54. Goldberg DP, Williams P. A user's guide to the General Health Questionnaire. Slough, UK: NFER-Nelson; 1988.
55. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: An introductory report. Psychol Med. 1983;13(3):595-605. https://doi.org/10.1017/s0033291700048017
56. UNESCO. International Standard Classification of Education (ISCED) 2011. Montreal, Canada: UNESCO Institute for Statistics; 2012.
57. Higgins JPT, Altman DG, Gøtzsche PC, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.
58. Snijders TAB, Bosker RJ. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. 2nd ed. Thousand Oaks, CA: SAGE Publishing; 2012.
59. Muthén LK, Muthén BO. Mplus User's Guide. 8th ed. Los Angeles, CA: Muthén & Muthén; 1998-2017.
60. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res. 1992;21(2):230-258. https://doi.org/10.1177/0049124192021002005
61. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling. 1999;6(1):1-55. https://doi.org/10.1080/10705519909540118
62. Kline RB. Principles and Practice of Structural Equation Modeling. New York, NY: Guilford Press; 2015.
63. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Methods. 2008;40(3): 879-891. https://doi.org/10.3758/brm.40.3.879
64. Miočević M, O'Rourke HP, MacKinnon DP, Brown HC. Statistical properties of four effect-size measures for mediation models. Behav Res Methods. 2018;50(1):285-301. https://doi.org/10.3758/s13428-017-0870-1
65. Schwarz G. Estimating the dimension of a model. Ann Stat. 1978;6(2):461-464. https://doi.org/10.1214/aos/1176344136
66. McLachlan GJ. On bootstrapping the likelihood ratio test statistic for the number of components in a normal mixture. Appl Stat. 1987;36(3):318-324. https://doi.org/10.2307/2347790
67. Bibriescas N, Whittaker TA. Detecting latent classes through mediation in regression mixture models. Struct Equ Modeling. 2023;30(3):449-457. https://doi.org/10.1080/10705511.2022.2137027
68. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using Mplus. Struct Equ Modeling. 2014;21(3):329-341. https://doi.org/10.1080/10705511.2014.915181
69. Burns GL, Patterson DR. Normative data on the Eyberg Child Behavior Inventory and Sutter-Eyberg Student Behavior Inventory: parent and teacher rating scales of disruptive behavior problems in children and adolescents. Child Fam Behav Ther. 2001;23(1): 15-28. https://doi.org/10.1300/J019v23n01_02

70. Leijten P, Thomaes S, de Castro BO, Dishion TJ, Matthys W. What good is labeling what's good? A field experimental investigation of parental labeled praise and child compliance. Behav Res Ther. 2016;87:134-141. https://doi.org/10.1016/j.brat.2016.09.008
71. Morris AS, Silk JS, Steinberg L, *et al*. The role of the family context in the development of emotion regulation. Soc Dev. 2007;16(2):361-388. https://doi.org/10.1111/j.1467-9507.2007.00389.x
72. Leijten P, Scott S, Landau S, *et al*. Individual participant data meta-analysis: impact of conduct problem severity, comorbid attention-deficit/hyperactivity disorder and emotional problems, and maternal depression on parenting program effects. J Am Acad Child Adolesc Psychiatry. 2020;59(8):933-943. https://doi.org/10.1016/j.jaac.2020.01.023
73. Kazdin AE. Research Design in Clinical Psychology. 6th ed. Cambridge, UK: Cambridge University Press; 2023.