The Earlier the Better? Individual Participant Data and Traditional Meta-analysis of Age Effects of Parenting Interventions

Gardner, F.; Leijten, P.; Melendez-Torres, G.J.; Landau, S.; Harris, V.; Mann, J.; Beecham, J.; Hutchings, J.; Scott, S.

DOI
10.1111/cdev.13138

Publication date
2019

Document Version
Final published version

Published in
Child Development

CHILD DEVELOPMENT

Child Development, January/February 2019, Volume 90, Number 1, Pages 7–19

The title for this Special Section is **Meta-analysis and Individual Participant Data Synthesis in Child Development**, edited by Glenn I. Roisman and Marinus H. van IJzendoorn

The Earlier the Better? Individual Participant Data and Traditional Meta-analysis of Age Effects of Parenting Interventions

Frances Gardner
University of Oxford

Patty Leijten
University of Amsterdam and University of Oxford

G.J. Melendez-Torres
Cardiff University

Sabine Landau and Victoria Harris
King's College London

Joanna Mann
University of Oxford

Jennifer Beecham
London School of Economics and Political Science

Judy Hutchings
Bangor University

Stephen Scott
King's College London

Strong arguments have been made for early intervention for child problems, stating that early is more effective than later, as the brain is more malleable, and costs are lower. However, there is scant evidence from trials to support this hypothesis, which we therefore tested in two well-powered, state-of-the-art meta-analyses with complementary strengths: (a) Individual participant data (IPD) meta-analysis of European trials of Incredible Years parenting intervention (k = 13, n = 1696; age = 2–11); (b) Larger, trial-level robust variance estimation meta-analysis of a wider range of parenting programs (k = 156, n = 13,378, M_{age} = 2–10) for reducing disruptive behavior. Both analyses found no evidence that intervention earlier in childhood was more effective; programs targeted at a narrower age range were no more effective than general ones.

Global policy directives are clear-cut in recommending early intervention (Allen, 2011; Black et al., 2017; WHO, 2016) for both mental and physical health problems, citing neuroscientific, economic and life course developmental research in support of these recommendations (Caspi et al., 2016; Heckman, 2006; Shonkoff & Fisher, 2013). But how strong is the evidence for a timing effect, whereby early interventions to prevent or reduce mental health difficulties are more effective than those delivered later in the child's life? A substantial body of evidence from behavioral and neuroscience suggests that children's development may be more malleable during the first few years of life, during periods of very rapid neural development (Wachs, Georgieff, Cusick, & McEwen, 2014). During these sensitive periods, the developing brain is thought to be more responsive to environmental influences, both those occurring naturally and those resulting from planned intervention. Arising from this body of the research is the critical question of timing:

This work was supported by grants from NIHR PHR 12-3070-04 to Gardner, Scott, Landau, Hutchings, Beecham, and the Swedish Board of Health and Welfare to Frances Gardner. Landau received salary support from NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. We thank all data site collaborators, and Wendy Knerr, University of Oxford, for her help with the systematic review for Meta-analysis 2.

Correspondence concerning this article should be addressed to Frances Gardner, Dept of Social Policy and Intervention, University of Oxford, 32, Wellington Sq, Oxford OX1 2ER., United Kingdom. Electronic mail may be sent to frances.gardner@spi.ox.ac.uk.

 0009-3920/2019/9001-0001
DOI: 10.1111/cdev.13138

when in children's development are interventions likely to have the strongest effect?

However, despite the theoretical attractions of intervening early when the brain is more plastic, there is very little empirical literature directly addressing this question. Heckman's (2006) work on timing of interventions made a strong call for investment in early intervention. He compared the effects of different interventions, from early childhood through to adolescence, and concluded there was substantially diminishing effectiveness and cost effectiveness with increasing age. However, these analyses have several important limitations. First, they compare different interventions at different ages, such as cognitive stimulation interventions in early childhood with delinquency reduction interventions in adolescence. Hence, they preclude like-for-like comparison of similar interventions and instead compare interventions at different ages that are likely to be very different in form, context and developmental mechanisms—all factors that may contribute to their effectiveness. Moreover, several interventions designed for older youth are known to be of limited effectiveness (e.g., boot camps, many employment schemes) and sometimes yield iatrogenic effects, for instance through peer contagion mechanisms (Dishion, McCord, & Poulin, 1999). A true test of the early intervention hypothesis requires a comparison of the effects of interventions that target plausible and similar underlying psychological mechanisms at different stages of child development.

Parenting Interventions

Parenting interventions provide an example of a well-established intervention, which can be implemented across a wide range of developmental stages, from infancy to late adolescence (Scott & Gardner, 2015). Parenting interventions aim to improve parent–child relationships and children's developmental outcomes, and have a substantial evidence base showing their effectiveness for reducing children's disruptive behavior (Leijten, Melendez-Torres, Knerr, & Gardner, 2016; Weisz et al., 2016). The majority of evidence-based parenting interventions are based on social learning theory. Such interventions include components on positive relationship building and discipline, for example, teaching warm, responsive play to parents, social reinforcement techniques, and proactive approaches to limit setting (Kaehler, Jacobs, & Jones, 2016; Leijten, Melendez-Torres, et al., 2018). Of course, there is much variation by developmental stage in expectations for children's behavior, and therefore in the form and focus of these parenting strategies. For example, as children start to play outside the home, new parenting skills for monitoring their whereabouts become salient (Dishion & McMahon, 1998; Shaw, Bell, & Gilliom, 2000) that may be different from those needed to monitor a toddler or a teenager. Importantly, despite these differences, social learning theory-based interventions target similar underlying parenting mechanisms, including positive behavioral support and clarity of expectations and reinforcers, combined with warmth and involvement (Leijten, Melendez-Torres, et al., 2018; Scott & Gardner, 2015). Because parenting interventions target similar mechanisms, using similar interventions, across a wide range of developmental stages, they are a good candidate for testing the hypothesis that early interventions are more effective than later ones.

Although the broad mechanisms tapped in social learning theory-based interventions appear to be similar across development, there nevertheless may be merit in implementing interventions that target narrower age ranges, as this affords the possibility of greater tailoring and specificity of the intervention content to that developmental stage. For example, Shaw et al. (2000) suggest that the transition to toddlerhood is a crucial developmental stage where parents may first encounter the need to deal with a mobile, defiant child, and interventions that help parents develop skills that are specific to this stage may be particularly effective. In addition to testing the early intervention hypothesis across developmental periods, there is a need for evidence as to whether interventions that focus on one specific child developmental period are more effective than interventions that span different developmental periods.

Evidence on the Early Intervention Hypothesis

There is surprisingly little direct empirical support for the early intervention hypothesis for parenting interventions. Systematic reviews are poorly set up to answer questions about age effects: many have not tested if early interventions are better; others have done so, but based on small samples of trials. This is because in conventional meta-regression, it is only possible to test the effects of age (or other moderators) at trial aggregate level. As a result, statistical power tends to be inadequate, because the sample size reflects the number of trials (not the number of families). Most reviews of randomized trials in the parenting field have included < 60 studies (e.g., Bakker, Greven, Buitelaar, & Glennon, 2017, $k = 17$; Comer, Chow, Chan, Cooper-Vince, & Wilson, 2013, $k = 36$; Dretzke et al., 2009, $k = 57$; Furlong et al., 2012, $k = 13$). Where sample sizes are larger, this often results from inclusion of nonrandomized designs, more likely to lead to biased estimates (e.g., Lundahl, Risser, & Lovejoy, 2006, $k = 63$). Where reviews have tested age effects, findings have been mixed, with some finding no age effects (e.g., Furlong et al., 2012; Lundahl et al., 2006) and others finding effects in either direction. Thus, a meta-analysis of 101 evaluations of Triple P parenting interventions, of which 74 were randomized (Sanders, Kirby, Tellegen, & Day, 2014), found that child outcomes improved to a greater extent in trials where the mean child age was younger, across the range 0–18 years (albeit children in most trials were young, $M_{age} = 5.9$ years). However, Comer et al.'s (2013) meta-analysis, which covered a wider range of parenting programs but a narrower age range (2–7), provided support for "later" rather than early interventions, finding greater effects on disruptive behavior in trials where the mean child age was older.

This Study

Meta-analysis 1: Individual Participant Data

A thorough examination of whether children's age influences the extent to which children benefit from parenting interventions requires a large sample of families with children from a wide age range. We therefore adopt an individual participant data (IPD) meta-analysis approach that synthesizes individual data from all families in a near-complete set of randomized trials of the same parenting intervention, the Incredible Years (IY), in Europe (Menting, Orobio de Castro, & Matthys, 2013; Webster-Stratton & Reid, 2010). We focused on this program for the following reasons: (a) It is a manualized intervention with a substantial evidence base (Gardner & Leijten, 2017; Menting et al., 2013), recommended by UK National Institute for Health and Care Excellence (NICE) and other policy bodies; (b) although it was developed in the USA, there has been widespread dissemination of IY in many European countries, across a range of ages 2–12; and (c) there are active European research networks for IY, raising the chances of obtaining data from a near-complete set of randomized trials for IPD meta-analysis from this region. We focused on Europe for the following additional reasons (a) most European trials have been conducted independently of the program developer—important because developer involvement is often associated with stronger intervention effects, and may represent a source of bias (Eisner, 2009). On the other hand, most US trials have been conducted by the program developer; (b) European countries that have implemented IY tend to have relatively similar health and social care systems (in contrast to the USA), which increases comparability of program effects across countries. Main effects of IY based on this IPD dataset were reported by Leijten, Gardner, et al. (2018).

IPD has substantial advantages over conventional meta-regression, which is limited to exploring between-trial variation in moderators such as age. This is because in a traditional review, the effects of age can only be coded at aggregate-level, for each trial (e.g., Lundahl et al., 2006; Sanders et al., 2014), resulting in loss of all information on within-trial variability in age and its relationship to outcome. Meta-analysis of trial data at the level of the individual participant solves this problem and brings several important advantages, including substantially raised power to test moderators, the ability to separate between- and within-trial moderation effects, and the opportunity to control for potential confounders of within-trial age effects, such as severity of behavior problems (Brown et al., 2013). By pooling IPD across trials and analyzing all data in the same way, it brings greater transparency and reduces potential bias (Riley, Lambert, & Abo-Zaid, 2010), an important consideration given mounting concern about bias in trials (Ioannidis et al., 2014) and the "replication crisis" in psychology. However, these transparency advantages only hold when investigators can access near-complete samples of trials for analysis.

Meta-analysis 2: Meta-Regression at Trial Level

Although IPD brings substantial advantages, its main drawback is limited generalizability, stemming from practical constraints. First, it is rare to obtain individual data from as many trials as is possible in aggregate level meta-analysis. Second, sample size may be further constrained by the fact that harmonizing data across trials, where trials have assessed similar concepts but used different measures, is very labor intensive.

We therefore aim to replicate our findings from Meta-analysis 1 in a conventional meta-analytic sample that includes many more trials and a wider range of parenting interventions and geographical and cultural contexts. Previous conventional meta-analyses are outdated and relatively small. We aimed to enhance power to detect age effects both by extensive literature searching and through state-of-the-art analytic techniques that harness information from multiple outcomes within each trial.

Together, the two meta-analyses will test three research questions. Our primary question is whether younger children benefit more than older children, by examining age as a continuous moderator. In addition, we address two related questions that are frequently raised but as yet unanswered. Second, can age effects be translated into children's developmental stages specifically? For example, are children more responsive to parenting interventions at particular developmental stages, such as the toddler and preschool years, compared to school age. Third, should interventions be developmentally specific? We test whether interventions that are targeted to a narrower age range (e.g., focused on one school year, e.g., school entry) are more effective than those targeted at a wider age range (e.g., 2–8 years). These two additional questions will be tested with the larger trial-level meta-analysis. By utilizing both IPD and conventional meta-analysis, we bring the twin strengths of each method to testing the primary question of whether earlier parenting interventions are more effective than ones delivered later in the child's life.

Meta-analysis 1, IPD

Method

Our IPD meta-analysis of IY programs in Europe follows PRISMA IPD reporting guidelines (Stewart et al., 2015). The study protocol is available at (blinded for review). Ethical approval was granted by (blinded for review).

Eligibility Criteria, Identifying, and Selecting Trials

We sought to include all data from all completed randomized trials of the IY parenting intervention in Europe, published or unpublished, for children aged 1–12 years, with no restriction on year of publication or included outcome measures (Appendix S1, flow chart). We excluded trials or conditions within trials that: (a) were not randomized; (b) included additional nonparenting programs, such as child-focused interventions; or (c) were highly abbreviated, nonstandard versions of the usual IY intervention of 12–14 sessions.

Trials were identified through: (a) systematic searches in five databases (CINAHL, Embase, Global Health, MEDLINE, PsycINFO) in January 2015; (b) the IY website library; (c) consultation with experts including European IY mentors' network. Searches via OVID used the following: (a) incredible years.mp; (b) webster-stratton.mp; (c) 1 or 2. Search strings were adapted for other databases. Eligibility was assessed by the first author and double checked by four additional authors, with no differences of opinion.

Data Collection and Data Integrity

All available fully anonymized data were requested for 15 identified trials of the IY parenting intervention (see Table 1). Five trials were not yet published at this time. Investigators signed data sharing agreements that specified ethical and ownership issues. One 2002 trial (#15) investigator had no longer retained the data. Raw, individual item-level data were supplied in SPSS for 14 trials and checked for missing items, scale validity and scores, internal consistency, baseline imbalance, and consistency with trial protocols and reports. Copies of original questionnaires were supplied to check for consistent use across trials. Any queries were resolved in collaboration with trial investigators. Risk of bias in trials was assessed with the Cochrane tool (Higgins & Green, 2011).

Data Items and Harmonization of Measures

Disruptive child behavior. We chose as the primary outcome measure for the meta-analysis, the Eyberg Child Behavior Inventory Intensity Scale (ECBI-I; Robinson, Eyberg, & Ross, 1980); this was used most frequently across trials (k = 11), assessed at baseline (before randomization) and postintervention. ECBI-I is a 36-item scale that assesses parent-reported frequency of disruptive behavior on a 7-point Likert scale and has demonstrated strong psychometric properties (Robinson et al., 1980). Two trials (#3; #14, n = 124; 141, respectively) used a different measure of disruptive behavior (Parental Account of Children's Symptoms [PACS]; Taylor, Schachar, Thorley, & Wieselberg, 1986), and in both cases, data were converted to a score on the ECBI-I, using norm deviation scores (Taylor, Sandberg, Thorley, & Giles, 1991, PACS; Robinson et al., 1980, ECBI). PACS and ECBI-I scores correlated r = .71 in our sample, based on data from four trials (#10–#13) that included both measures. Internal consistency at baseline was high (ECBI-I α = .94; PACS, α = .82). Data on the parent who was the primary caregiver (98% mothers) were used because few trials include data on conduct problems reported by both parents. There were very limited data (k = 3) available on an alternative measure of conduct problems, by teacher report, hence these were excluded.

Table 1
Characteristics of the 15 Trials That Met Inclusion Criteria, Meta-analysis 1, Individual Participant Data (IPD)

Trial	Author (year)	Country	Setting	Conduct problems?	*N*	Child age (*M*)	% Poverty	% Ethnic
#1	Larsson et al. (2009)	Norway	Clinics	Yes	75	3–8 (6.58)	25	1
#2	Axberg and Broberg (2012)	Sweden	Clinics	Yes	62	3–8 (5.97)	41	0
#3	Seabra-Santos et al. (2016)	Portugal	University clinics	Yes	124	3–6 (4.66)	0	0
#4	McGilloway et al. (2012)	Ireland	Community	Yes	149	2–7 (4.84)	47	6
#5	Menting, de Castro, Wijngaards-de Meij, and Matthys (2014)	Netherlands	Community	No	99	1–11 (6.30)	93	78
#6	Leijten, Raaijmakers, Orobio de Castro, van den Ban, and Matthys (2017)	Netherlands	Clinics, schools	Partly	156	2–8 (5.59)	74	65
#7	Hutchings et al. (2007)	Wales	Community	Yes	153	3–4 (3.84)	79	1
#8	Hutchings, Griffith, Bywater, and Williams (2017)	Wales	Community	No	103	0–2 (1.85)	56	0
#9	Morpeth et al. (2017)	England	Community	Yes	161	2–4 (3.68)	63	52
#10	Scott, Sylva, et al. (2010)	England	Schools	Yes	112	4–6 (5.21)	44	40
#11	Scott, O’Connor, et al. (2010)	England	Schools	No	174	4–6 (5.50)	44	75
#12	Scott, Sylva, Kallitsoglou, and Ford (2014)	England	Schools	Yes	214	3–7 (6.07)	80	19
#13	Gardner, Burton, and Klimes (2006)	England	Community	Yes	76	2–9 (5.93)	64	2
#14	Scott, Spender, Doolan, Jacobs, and Aspland (2001)	England	Clinics	Yes	141	2–10 (5.67)	58	15
#15	Patterson et al. (2002)	England	General Practice	Yes	116	2–8	25	9

Note. #15 = IPD not available; #8 = IPD supplied, but no data on primary outcome.

Child age. Age was coded for each child as a continuous variable, in months.

Statistical Methods

Power calculations for an anticipated sample size of $N = 1{,}400$ participants gave 97% power to detect a small interaction effect between two binary variables ($f = 0.1$) using an analysis of variance F-test at the 5% significance level. Formal analyses, conducted in Stata v.14, (StataCorp LLC, College Station, TX, USA) used the pooled dataset harmonized from 13 trials; a 14th (#8) had no data on the primary outcome, as children were aged 12–24 months. The purpose of the statistical analyses was to assess whether baseline child age moderated the effect of IY on disruptive behavior (ECBI-I) postintervention. Three statistical issues needed addressing: (a) the pooled data had a hierarchical structure with families (Level 1) nested within parenting groups (Level 2) within the intervention arm, and parenting groups nested within trials (Level 3); (b) there was some variation in design features of the original trials that needed accounting for, such as stratified randomization, and changes in allocation ratios over the trial duration; (c) it was necessary to minimize any missing data biases. We addressed these issues using a one-stage model that, in one step, models the IPD to answer the moderation questions. One stage models carry the advantage of greater efficiency in terms of power when between-trial and within-trial moderation effects do not differ (Fisher, Carpenter, Morris, Freeman, & Tierney, 2017).

We used multilevel/mixed effects modeling with post ECBI-I as the dependent variable, with fixed effects for trial arm, trial mean age (between-trial variable), participant age deviation from trial mean age (within-trial variable), and respective interaction terms. Tests of the effects of the interaction terms then provided assessments of the between-trial and within-trial moderating effects of age. Importantly this allowed us to assess empirically whether these two moderating effects differed; if such a difference was statistically significant at a liberal 10% test level then two separate moderating effects were allowed, if not a more powerful model with a single interaction term was fitted. The size of any IY effect moderation was described by a moderation index, which was constructed as the change in IY effect on posttest ECBI-I per one (pooled sample) standard deviation change in baseline age.

The hierarchical structure of the data was modeled by random intercepts that varied with parenting group within the active arm of a trial (Level 2) and a further random intercept that varied with trial (Level

3). Trial design features were accounted for by including relevant fixed effects (e.g., for randomization stratifiers) or random intercepts that varied with cluster in a cluster randomized trial. Known predictors of post-test ECBI-I (baseline ECBI-I and child gender) were included as fixed effects, as was the possible confounder of prevention versus treatment trial, and its interaction with trial arm, in order to adjust moderation effects. Finally, in order to allow for further treatment effect heterogeneity, a trial-varying random coefficient of trial arm was included in the model.

The IPD was subject to missing values in moderator and outcome variables. In order to produce valid estimates of moderation effects under a missing at random assumption we used multiple imputation, specifically the multiple imputation by chained equations approach (White, Royston, & Wood, 2011).

Results

Study Characteristics

Fifteen IY trials met inclusion criteria (Table 1), conducted in England ($k = 7$), Wales ($k = 2$), Netherlands ($k = 2$), with one each in Ireland, Norway, Portugal, and Sweden. However, two (7%) UK trials were excluded from the meta-analysis, one because the first author (trial #15), reported that data were no longer available; another trial (#8) supplied IPD but no data on the primary outcome, as children were aged 10–24 months. Thus, 13 trials ($N = 1,696$) were included in the analyses reported here. Due to uneven randomization ratios in some trials, there were 1,046 in the intervention arm and 650 in the control arm. For all trials, we included data for baseline and the first postintervention assessment, which was in most cases 3–6 months later, or 1–2 months after the end of 12- to 14-week intervention; in most studies this was the end-point of the randomized part of the evaluation. Risk of bias within studies was assessed as low on most items; across studies it was also low with regard to availability of IPD, as all but one eligible trial supplied data.

Ten trials were treatment trials (defined by referral for high levels of conduct problems, to specialist services), or indicated prevention trials (children screened for high levels of disruptive behavior). Three were selective prevention trials (targeting high-risk families, e.g., socioeconomically disadvantaged families; mothers released from prison). Overall, most trials ($k = 10$) included families who were predominantly socially disadvantaged, by having low income or a lone parent. Table 2 shows the demographic characteristics across trials, indicating that a majority of families (58%) had low income, and 30% were from ethnic minorities. Six trials in urban areas of the UK and Netherlands accounted for over 90% of the families from ethnic minorities (range 19%–78% per trial). The mean age of children

Table 2
Summaries for Demographics and Clinical Outcome by Randomized Group, Pooled Sample, 13 Trials

	Total *N* children, max 1,696	No. of trials info available	Control (max *N*, 650)		Incredible Years (max *N*, 1,046)	
	N	*k*	*N*	%	*N*	%
Categorical variables						
Child gender (male)	1,696	13	650	63.8	1,046	63.1
Low income	1,614	13	615	57.9	999	57.6
Low education	1,696	13	650	35.5	1,046	40.5
Lone parent	1,606	13	606	33.0	1,000	36.8
Teen parent	1,609	12	605	12.6	1,004	11.7
Unemployed	1,303	11	522	30.3	781	37.5
Ethnic minority	1,651	13	629	30.0	1,022	30.9
Child conduct problem score ECBI-I	*N*	*k*	*N*	*M* (*SD*)	*N*	*M* (*SD*)
Continuous variables						
Baseline	1,622	13	611	135.5 (37.0)	1,011	139.4 (37.0)
Postintervention	1,445	13	567	125.5 (37.9)	878	116.2 (34.7)
Child age, months	1,682	13	643	64.2 (16.9)	1,039	62.4 (18.3)

Note. ECBI-I = Eyberg Child Behavior Inventory Intensity Scale.

was 5 years (range 22–10); one quarter of parents reported significant levels of depression. In nine trials, the control condition was a wait list, who were offered IY 6–12 months later; in two trials there was a minimal intervention and two no intervention.

Main Effect of the Intervention

There was a significant overall effect of the intervention ($z = 10.08$, $p < .001$), reported in Leijten, Gardner, et al. (2018), estimated to be a reduction of 13.5 points (95% CI [10.9, 16.1]) on the ECBI-I. Most trials found that the IY intervention reduced child conduct problems, with standardized effect sizes varying across trials from very small (−0.12) to large (−0.76), with overall a moderate effect size (−0.43). There was moderate between-trial heterogeneity in program effects ($I^2 = 42.5\%$.) For trial #15, the published findings reported no effect sizes but showed significant effects on one of two measures of conduct problems at post test, and both measures at 6 month follow up.

Moderation by Age

We found no evidence that any IY effect moderation by age varied between the trial and individual level ($p = .45$), nor was there any suggestion that the relationship between post-test ECBI and age was not linear ($p = .89$). We therefore employed a parsimonious linear model with a single interaction effect for age. After adjusting for baseline ECBI-I and gender this moderation effect was very small in size (a modification index of 0.04 points on the ECBI-I scale, which translates into a standardized regression coefficient of $0.04/31.4 = 0.0013$ on a correlation scale) and not statistically significant according to a formal test ($p = .65$; 95% CI [−0.1, 0.2] points). There was therefore no evidence to suggest that child age moderated the benefit of the IY intervention.

Meta-analysis 2, Trial-Level

Method

Sources, Study Selection, Inclusion Criteria, Data Extraction

We aimed to test age effects of parenting interventions, by conducting a meta-analysis at trial aggregate level, in a larger and more diverse sample of trials of parenting interventions than prior reviews or than is possible with IPD. We identified randomized trials of parenting interventions for reducing disruptive child behavior that taught parents skills based on social learning theory perspectives. We updated our systematic literature search from Leijten et al., 2016, using six online databases (e.g., MEDLINE), to include studies up to January 2016 (see Appendix for characteristics of trials, Appendix S3, 4 for search and search update strategies, Appendix S5 for flow chart). To maximize the number of relevant trials for analysis, we also searched for unpublished studies in trial registries and by contacting experts in many countries. Trials were not excluded based on date or language. Inclusion criteria were as follows: (a) comparing a parenting intervention based on social learning theory principles to any type of control condition; (b) random assignment to conditions; (c) more than 50% of intervention sessions focused on parenting; and (d) children's mean age at trial level was between 2 and 10 years. We excluded interventions for parents of special populations such as children in foster care or with disabilities. One researcher assessed abstracts and full texts of studies that were likely to meet inclusion criteria. Uncertainties and the final list of studies included in the review were assessed by the first and third author. We extracted the following data: mean child age of sample (in years), range of child age in sample (expressed as number of years between the oldest and youngest child in the sample), developmental stage(s) of the children included in the trial (toddler, preschool, lower primary, upper primary, or combinations of these).

Effect size calculation. We converted effect sizes into Cohen's d values based on within-trial arm means and standard deviations reported at post-treatment. As recommended in the analysis of randomized trials, we preferred estimates of trial arm differences that were analysis of covariance adjusted for baseline. Where needed, we used alternative summary statistics (e.g., p-values and sample sizes, or t-test statistics) to calculate Cohen's d values.

Risk of bias. We assessed the risk of bias in each study (as high, low, or unclear) using the Cochrane Collaboration tool (Higgins & Green, 2011).

Analytic Strategy

Most studies included multiple measures of disruptive child behavior, and hence multiple effect sizes. Various approaches to address this challenge exist, including selection-based protocols (i.e., decision rules to select the "most appropriate" effect size), multivariate meta-analysis, and robust variance estimation approaches (Tanner-Smith & Tipton, 2014). For testing the moderating effect of child age, expressed as a trial-level summary, we chose a robust variance estimation approach, as selection-based protocols are prone to bias and lose information from included studies, and multivariate meta-analysis is appropriate when effect sizes are correlated but target different outcome concepts. Robust variance estimation meta-analysis reweights the multiple effect sizes within studies using an approximate variance–covariance matrix, resulting in valid point estimates and significance tests even when the exact variance–covariance matrix of effect sizes within studies remains unknown (Hedges, Tipton, & Johnson, 2010). All analyses were estimated assuming an intercorrelation within studies of $\rho = .8$ and random effects. In these models, a negative effect size is indicative of greater effectiveness; thus, a positive coefficient is interpreted as a decrease in effectiveness.

Because this is a meta-regression, we labeled trials as to the mean age in the sample. To account for any phase effects, we also categorized trials into one of three groups depending on mean age: toddlers and preschool (ages 1–6), school age (ages 6–12) and combined (ages 1–12). We also coded range of age as a continuous variable. We explored mean age, age group, and age range in different univariate meta-regressions, and then estimated exploratory models including both age and age range, and including interactions between age and age range.

Results

Included Studies

We found 154 trials meeting inclusion criteria for our robust variance meta-analysis (388 effect sizes, 13,387 participants). Table 3 and Appendix S2 show the diverse range of trials, which include 50 different parenting programs from 22 countries, with *N*s ranging from 17 to 695. The average effect size of the parenting interventions on disruptive child behavior was $d = -.47$ (95% CI [−0.55, −0.40]). Mean child age at trial level was 5.3 years, *SD* 1.8, range 2–10.

First, our primary question of whether younger children benefit more than older children: We found no evidence of any moderation effect by age ($\beta = .016$; 95% CI [−0.029, 0.062]), in other words, the effectiveness of the intervention for reducing child disruptive behavior did not vary by the average age of the children in the trial. Relatedly, developmental stage did not moderate outcome (school age: $\beta = .05$; 95% CI [−0.11, 0.21]; all ages: $\beta = .26$, 95% CI [−0.07, 0.59]; preschool age as reference category). Thus interventions were no more effective in the preschool than in the school age era. Second, the question of whether interventions should be developmentally specific: We found no evidence of moderation by age range ($\beta = -.01$; 95% CI [−0.04, 0.02]), thus children involved in interventions targeting a narrower range of ages did not fare any better than those in interventions targeting a wider range of ages. Interaction models did not yield any significant effects and did not change interpretation of univariate meta-regression findings.

Table 3
Summary of Studies Included in Meta-analysis 2

	$k = 154$ $n = 388$
Sample	
Total number of children	13,387
Child age range (*M*)	2–10 (4.93)
Program (%)	
Triple P	33
Incredible Years	24
Parent–child interaction therapy	9
Other	34
Geographical region (%)	
North America	36
Northwest Europe	27
Australia/New Zealand	27
Other	10

Note. k = number of studies; n = number of effect sizes.

General Discussion

Using two state-of-the-art methods of meta-analysis with important and complementary advantages over conventional approaches, we find no evidence for any influence of younger child age on the effectiveness of parenting programs for improving children's disruptive behavior. Hence there was no support for the early intervention hypothesis. Our IPD meta-analytic findings show that in trials of the IY parenting program, across multiple countries, conducted by different teams all independent of the program developer, child behavior is equally open to change at older as younger ages, across the range 2–11 years. The robust variance estimation meta-analysis replicated the IPD finding in a more diverse sample of trials, the largest meta-analytic sample to date in this field, with a wide range of parenting interventions based on social learning theory. When translating age effects into developmental stages, we found no difference in effects in the toddler and preschool phases compared to school age. This meta-analysis also tested if parenting interventions were more effective when targeted at a narrower age range and therefore able to be better tailored to a particular developmental phase. We found no added benefit of these potentially more developmental stage-specific programs.

Why did our findings not support the dominant early intervention hypothesis? There are several potential explanations. First, it may be that the plasticity of child behavior in response to changes in parenting is similar across childhood years. This would be consistent with social learning theory explanations, whereby coercive cycles of parent–child interaction contribute substantially to child disruptive behavior at all ages (Patterson, 1982). If so, then changing parenting in ways that reduces these cycles may have a similar impact at all ages.

Although coercion theory is not developmentally specific, it suggests that patterns of parent–child interaction become more entrenched over time, and thus harder to change. Although our study did not measure age of onset of conduct problems, we could speculate that older children in our study may tend (on average) to have had longer experience of family coercion. Nevertheless, our findings do not support the notion that these potentially more entrenched parent–child interactions are harder to change.

A second possible explanation is that underneath seemingly similar levels of disruptive behavior in younger and older children may lie different subtypes of disruptive behavior, which in turn influence how malleable problems are. If very young children show severe disruptive behavior, this might reflect the presence of "early onset type" problems, which may be more likely to have neurobiological origins (Caspi et al., 2016), and to predict greater persistence and ultimately severity of antisocial behavior. It is possible that this factor offsets any malleability benefit at younger ages. The data unfortunately do not allow further test of this explanation, as we do not know how many older children also had early onset behavior problems.

Why were studies targeting more specific developmental stages not more effective? This can be explained in similar ways to the lack of age effect. If parenting mechanisms thought to influence children's behavioral development tend to be similar across ages (e.g., coercion, warmth, joint involvement, positive behavioral support), then highly developmentally targeted programs are not needed. However, because these key mechanisms are expressed differently depending on the child's age, then, as is common in many programs, to optimize effectiveness, delivery staff should be well trained to adapt the content to children's individual needs and developmental stage (Gardner & Leijten, 2017).

Our findings overall have a number of policy implications. First, although it is vital not to delay intervention, so as to minimize the period of upset and suffering caused by disruptive behavior, these findings are optimistic in that it is not in any sense "too late" to intervene later in childhood, when children are older. Second, they point to the need to ensure services focus on identifying and supporting older and younger children with evidence-based parenting interventions rather than focusing a disproportionate share of intervention resources toward younger children. This is underlined by our pooled IPD economic analyses for a UK subsample of the 13 trials (k = 5, n = 608), which found that IY is likely to be *more* cost-effective for children older, rather than younger than 5 years of age (Gardner et al., 2017). Thus, for evidence-based parenting interventions, our overall policy message on effectiveness and cost effectiveness ("never too early, never too late") contrasts with that from Heckman's (2006) well-known economic analysis ("the earlier the better"). Third, our findings suggest there may not be a need for different programs for specific developmental stages, so long as they are sensitively adapted to the age of the particular child. This would have significant implications for services in relation to cost saving, both in terms of therapist training and also intervention delivery. That the same parenting interventions can be effective for children from toddlerhood to middle childhood is an important argument against a tendency to increasing age specificity of programs. This is echoed in the findings of other meta-analyses, which also find no evidence pointing to a need for greater specificity of interventions, for example, for different cultural groups. Thus, recent work has found similar effect sizes across disparate countries and cultures (Gardner, Knerr, & Montgomery, 2016; Leijten et al., 2016), and, in IPD meta-analysis, across ethnic and social groups (Gardner et al., 2018). Finally, although our findings pertain only to parenting interventions, potentially they have wider implications for early intervention policy. They remind us that, in the absence of adequately powered meta-analyses of randomized trials (preferably employing individual-level data), we cannot assume that for child development interventions earlier is necessarily better.

We draw attention to several limitations of the studies. Both meta-analyses, although covering a wide age range, were limited to childhood, between ages 2 and 11. We do not know whether very early parenting interventions from ages 0 to 2 are any more or less effective than those delivered later. Nor were we able to test whether the early intervention hypothesis might hold for childhood versus adolescence or adulthood, or indeed if there may be further sensitive periods when children are more malleable in adolescence (Wachs et al., 2014). Indeed, many well-conducted independent replications of parent-focused interventions for disruptive behavior in adolescence have failed to show effectiveness, for example, the UK trials of Functional Family Therapy (Humayun et al., 2017) and Multisystemic Therapy (Fonagy et al., 2018). Clearly there is a need to investigate age effects in other developmental periods. Second, both studies relied on parent-reported outcomes of intervention effects, which may be open to bias. However, there is evidence that effect sizes for directly observed child behavior outcomes are comparable to those for parent report (Menting et al., 2013; van Aar, Leijten, Orobio de Castro, & Overbeek, 2017). Third, we were only able to examine the effects of parenting programs on disruptive child behavior specifically rather than other child outcomes that may benefit from parenting interventions, such as emotional problems or cognitive development. It might be that the early intervention hypothesis holds for some outcomes and not for others. However, disruptive behavior predicts many marked impairments later in life and is the commonest problem in childhood, so it is not an insignificant issue. Fourth, our review concerned programs that were social learning theory based and therefore cannot tell us whether parenting programs based on changing other aspects of parenting may show an early intervention effect. For example, it may be that some aspects of parent–child interaction (e.g., attachment quality) develop during a sensitive period and are harder to repair later. Although our findings may not apply to other interventions, they are nevertheless very significant, as these parenting interventions have been widely disseminated in many countries and have probably the most extensive evidence base of any childhood psychosocial intervention. Fifth, few of the trials had sufficient long-term data for analysis, hence we cannot tell if the lack of age effects found here would be mirrored in longer follow-up data. However, in recent (albeit much smaller) aggregate-level meta-analysis of longer term effects of parenting interventions, van Aar et al. (2017) similarly found no evidence of moderation by age across the range 1–10 years.

There are limitations of each meta-analysis: IPD meta-analysis necessarily makes a number of assumptions in harmonizing data across trials (Brown et al., 2013). Although this study only evaluated one program, it included families from a diverse range of settings, countries, and ethnicities. Regarding Meta-analysis 2, it should be noted that our analyses of developmental specificity of interventions are based on data on the age range of the children in the study and not on how explicit or accurate was the intervention in its developmental targeting. However, in most cases, the age of the children in the trial reflected the range for which the program was intended. Including only a narrow age range allows program developers to design content that is more developmentally specific and makes the job of tailoring to individuals simpler for those delivering the program. Nevertheless, targeting a narrower age range did not predict better outcomes.

Our studies make a unique contribution to the use of meta-analysis in the developmental domain by testing age effects cumulatively, using two complementary approaches to meta-analysis. IPD is exceptionally well-powered, benefiting from fully utilizing all information about within-trial variation in age. By re-analyzing an unusually complete set of independent trials of the same program in Europe, we reduce risk of reporting bias and false positive results. Aggregate meta-analysis has few of these advantages but instead brings greater generalizability by permitting synthesis of many more trials, and examining whether developmentally targeted interventions are more effective than those serving wider age ranges. To our knowledge, this meta-analysis is larger and more up to date than other syntheses of randomized trials in childhood in this field. Together with the use of robust variance estimation, which takes advantage of all available outcome information on disruptive behavior outcomes, our study is likely to be better powered than other aggregate level meta-analyses for testing age as a moderator. Thus it provides a vital, potentially more generalizable complement to the still greater power and precision of our unique IPD meta-analysis. Importantly, both methods pointing to the conclusion that the abilities of the IY program, and other parenting interventions based on social learning theory to reduce disruptive child behavior are unaffected by the age and developmental stage of the child.

References

Allen, G. (2011). *Early intervention: The next steps*. Independent report to Her Majesty's Government. London, UK: HM Government.

Axberg, U., & Broberg, A. G. (2012). Evaluation of "the Incredible Years" in Sweden: The transferability of an American parent-training program to Sweden. *Scandinavian Journal of Psychology*, *53*, 224–232.

Bakker, M., Greven, C., Buitelaar, J., & Glennon, J. (2017). Practitioner review: Psychological treatments for children and adolescents with conduct disorder: Systematic review and meta-analysis. *Journal of Child Psychology & Psychiatry*, *58*, 4–18.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., . . . Devercelli, A. E. (2017). Early childhood development coming of age: Science through the life course. *The Lancet*, *389*, 77–90.

Brown, C. H., Sloboda, Z., Faggiano, F., Teasdale, B., Keller, F., Burkhart, G., . . . Muthén, B. (2013). Methods for synthesizing findings on moderation effects across multiple randomized trials. *Prevention Science*, *14*, 144–156.

Caspi, A., Houts, R. M., Belsky, D. W., Harrington, H., Hogan, S., Ramrakha, S., . . . Moffitt, T. E. (2016). Childhood forecasting of a small segment of the population with large economic burden. *Nature Human Behaviour*, *1*, 0005.

Comer, J. S., Chow, C., Chan, P. T., Cooper-Vince, C., & Wilson, L. A. (2013). Psychosocial treatment efficacy for disruptive behavior problems in very young children: A meta-analytic examination. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*, 26–36.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, *54*, 755.

Dishion, T. J., & McMahon, R. J. (1998). Parental monitoring and the prevention of child and adolescent problem behavior: A conceptual and empirical formulation. *Clinical Child and Family Psychology Review*, *1*, 61–75.

Dretzke, J., Davenport, C., Frew, E., Barlow, J., Stewart-Brown, S., Bayliss, S., . . . Hyde, C. (2009). The clinical effectiveness of different parenting programmes for children with conduct problems: A systematic review of randomised controlled trials. *Child & Adolescent Psychiatry & Mental Health*, *3*, 7.

Eisner, M. (2009). No effects in independent prevention trials: Can we reject the cynical view? *Journal of Experimental Criminology*, *5*, 163–183.

Fisher, D. J., Carpenter, J. R., Morris, T. P., Freeman, S. C., & Tierney, J. F. (2017). Meta-analytical methods to identify who benefits most from treatments: Daft, deluded, or deft approach? *British Medical Journal*, *356*, j573.

Fonagy, P., Butler, M., Cottrell, D., Scott, S., Pilling, S., Eisler, I., . . . Goodyer, I. (2018). Multisystemic therapy versus management as usual in the treatment of adolescent antisocial behaviour (START): A randomised controlled pragmatic effectiveness superiority trial. *The Lancet Psychiatry*, *5*, 119–133.

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.cd008225.pub2

Gardner, F., Burton, J., & Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: Outcomes and mechanisms of change. *Journal of Child Psychology & Psychiatry*, *47*, 1123–1132.

Gardner, F., Knerr, W., & Montgomery, P. (2016). Transporting evidence-based parenting programs for child problem behavior (age 3–10) between countries: Systematic review and meta-analysis. *Journal of Clinical Child and Adolescent Psychology*, *45*, 749–762.

Gardner, F., & Leijten, P. (2017). Incredible Years parenting interventions: Current effectiveness research and future directions. *Current Opinion in Psychology*, https://doi.org/10.1016/j.copsyc.2017.02.023

Gardner, F., Leijten, P., Mann, J., Landau, S., Harris, V., Beecham, J., Bonin, E. M., Hutchings, J., & Scott, S. (2017). Could scale-up of parenting programmes improve child disruptive behaviour and reduce social inequalities? Using individual participant data meta-analysis to establish for whom programmes are effective and cost-effective. Southampton (UK): NIHR Journals Library; 2017 Dec. (Public Health Research, No. 5.10.) Available from: https://www.ncbi.nlm.nih.gov/books/NBK469806/ doi: 10.3310/phr05100

Gardner, F., Leijten, P., Harris, V., Mann, J., Scott, S., & Landau, S. (2018). Do parenting interventions increase social inequalities in child conduct problems? Individual participant data (IPD) meta-analysis. In submission.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*, 1900–1902. https://doi.org/10.1126/science.1128898

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65.

Higgins, J., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0). Chichester, UK: Cochrane Collaboration, Wiley.

Humayun, S., Herlitz, L., Chesnokov, M., Doolan, M., Landau, S., & Scott, S. (2017). Randomized controlled trial of Functional Family Therapy for offending and antisocial behavior in UK youth. *Journal of Child Psychology & Psychiatry*. https://doi.org/10.1111/jcpp.12743

Hutchings, J., Gardner, F., Bywater, T., Daley, D., Whitaker, C., Jones, K., . . . Edwards, R. T. (2007). Parenting intervention in sure start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *British Medical Journal*, *334*, 678.

Hutchings, J., Griffith, N., Bywater, T., & Williams, M. E. (2017). Evaluating the Incredible Years toddler parenting programme with parents of toddlers in disadvantaged (flying start) areas of Wales. *Child: Care, Health and Development*, *43*, 104–113. (IPD supplied, but no data on disruptive child behaviour)

Ioannidis, J. P., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., . . . Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, *383*, 166–175.

Kaehler, L. A., Jacobs, M., & Jones, D. J. (2016). Distilling common history and practice elements to inform dissemination: Hanf-model BPT programs as an example. *Clinical Child and Family Psychology Review*, *19*, 236–258. https://doi.org/10.1007/s10567-016-0210-5

Larsson, B., Fossum, S., Clifford, G., Drugli, M. B., Handegård, B. H., & Mørch, W. T. (2009). Treatment of oppositional defiant and conduct problems in young Norwegian children. *European Child & Adolescent Psychiatry*, *18*, 42–52.

Leijten, P., Gardner, F., Landau, S., Harris, V., Mann, J., Hutchings, J., . . . Scott, S. (2018). Harnessing the power of individual participant data in a meta-analysis of the benefits and harms of the Incredible Years parenting program. *Journal of Child Psychology and Psychiatry*, *59*, 99–109.

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

Leijten, P., Melendez-Torres, G. J., Knerr, W., & Gardner, F. (2016). Transported versus homegrown parenting interventions for reducing disruptive child behavior: A multilevel meta-regression study. *Journal of the American Academy of Child & Adolescent Psychiatry.*, *55*, 610–617. https://doi.org/10.1016/j.jaac.2016.05.003

Leijten, P., Raaijmakers, M. A., Orobio de Castro, B., van den Ban, E., & Matthys, W. (2017). Effectiveness of the Incredible Years parenting program for families with socioeconomically disadvantaged and ethnic minority backgrounds. *Journal of Clinical Child & Adolescent Psychology*, *46*, 59–73.

Lundahl, B., Risser, H., & Lovejoy, M. C. (2006). A meta-analysis of parent training: Moderator and follow up effects. *Clinical Psychology Review*, *26*, 86–104. https://doi.org/10.1016/j.cpr.2005.07.004

McGilloway, S., Ni Mhaille, G., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., . . . Donnelly, M. (2012). A parenting intervention for childhood behavioral problems: A randomized controlled trial in disadvantaged community-based settings. *Journal of Consulting & Clinical Psychology*, *80*, 116–127.

Menting, A. T., de Castro, B. O., Wijngaards-de Meij, L. D., & Matthys, W. (2014). A trial of parent training for mothers being released from incarceration and their children. *Journal of Clinical Child & Adolescent Psychology*, *43*, 381–396.

Menting, A. T. A., Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clinical Psychology Review*, *33*, 901–913. https://doi.org/10.1016/j.cpr.2013.07.006

Morpeth, L., Blower, S., Tobin, K., Taylor, R. S., Bywater, T., Edwards, R. T., . . . Berry, V. (2017). The effectiveness of the Incredible Years pre-school parenting programme in the United Kingdom: A pragmatic randomised controlled trial. *Child Care in Practice*, *23*, 141–161.

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia.

Patterson, J., Barlow, J., Mockford, C., Klimes, I., Pyper, C., & Stewart-Brown, S. (2002). Improving mental health through parenting programmes: Block randomised controlled trial. *Archives of Disease in Childhood*, *87*, 472–477. (IPD data not available)

Riley, R. D., Lambert, P. C., & Abo-Zaid, G. (2010). Meta-analysis of individual participant data: Rationale, conduct, and reporting. *British Medical Journal*, *340*, c221.

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child & Adolescent Psychology*, *9*, 22–28.

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 Psychology Review*, *34*, 337–357.

Scott, S., & Gardner, F. (2015). Parenting programmes. In A. Thapar, D. Pine, J. Leckman, S. Scott, M. Snowling, & E. Taylor (Eds.), *Rutter's child and adolescent psychiatry* (6th ed., pp. 483–495). Chichester, UK: Wiley Blackwell.

Scott, S., O'Connor, T. G., Futh, A., Matias, C., Price, J., & Doolan, M. (2010). Impact of a parenting program in a high-risk, multi-ethnic community: The PALS trial. *Journal of Child Psychology & Psychiatry*, *51*, 1331–1341.

Scott, S., Spender, Q., Doolan, M., Jacobs, B., & Aspland, H. (2001). Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *British Medical Journal*, *323*, 194–198.

Scott, S., Sylva, K., Doolan, M., Price, J., Jacobs, B., Crook, C., & Landau, S. (2010). Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: The SPOKES project. *Journal of Child Psychology & Psychiatry*, *51*, 48–57.

Scott, S., Sylva, K., Kallitsoglou, A., & Ford, T. (2014). *Which type of parenting programme best improves child behaviour and reading? Follow-up of the helping children achieve trial*. London, UK: Nuffield Foundation.

Seabra-Santos, M. J., Gaspar, M. F., Azevedo, A. F., Homem, T. C., Guerra, J., Martins, V., . . . Moura-Ramos, M. (2016). Incredible Years parent training: What changes, for whom, how, for how long? *Journal of Applied Developmental Psychology*, *44*, 93–104.

Shaw, D. S., Bell, R. Q., & Gilliom, M. (2000). A truly early starter model of antisocial behavior revisited. *Clinical Child and Family Psychology Review*, *3*, 155–172.

Shonkoff, J. P., & Fisher, P. A. (2013). Rethinking evidence-based practice and two-generation programs to create the future of early childhood policy. *Development and Psychopathology*, *25*, 1635–1653.

Stewart, L. A., Clarke, M., Rovers, M., Riley, R., Simmonds, M., Stewart, G., & Tierney, J. F. (2015). Preferred reporting items for a systematic review and meta-analysis of individual participant data: The PRISMA-IPD statement. *Journal of the American Medical Association*, *313*, 1657–1665.

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*, 13–30. https://doi.org/10.1002/jrsm.1091

Taylor, E., Sandberg, S., Thorley, G., & Giles, S. (1991). *The epidemiology of childhood hyperactivity (Maudsley Monograph No 33)*. Oxford, UK: Oxford University Press.

Taylor, E., Schachar, R., Thorley, G., & Wieselberg, M. (1986). Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *British Journal of Psychiatry*, *149*, 760–767.

van Aar, J., Leijten, P., Orobio de Castro, B., & Overbeek, G. (2017). Sustained, fade-out or sleeper effects? A multilevel meta-analysis of parenting interventions for disruptive child behavior. *Clinical Psychology Review*, *51*, 153–163.

Wachs, T. D., Georgieff, M., Cusick, S., & McEwen, B. S. (2014). Issues in the timing of integrated early interventions: Contributions from nutrition, neuroscience, and psychological research. *Annals of the New York Academy of Sciences*, *1308*, 89–106.

Webster-Stratton, C., & Reid, M. J. (2010). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for children with conduct disorders. In J. R. Weisz & A. E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (pp. 194–210). New York, NY: Guilford.

Weisz, J. R., Kuppens, S., Ng, M. Y., Eckshtain, D., Ugueto, A. M., Vaughn-Coaxum, R., . . . Weersing, V. R. (2016). What five decades of research tells us about the effects of youth psychological therapy: A multilevel meta-analysis and implications for science and practice. *American Psychologist*, *72*, 79–117.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*, 377–399.

WHO. (2016). *INSPIRE: Seven strategies for ending violence against children*. Geneva, Switzerland: WHO.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's website:

Appendix S1. PRISMA Flowchart, Meta-Analysis 1

Appendix S2. Characteristics of Included Trials, Meta-Analysis 2

Appendix S3. Search Strategy Meta-Analysis 2

Appendix S4. Systematic Reviews and Reviews of Reviews Searched to Identify Relevant Studies for Meta-Analysis 2

Appendix S5. PRISMA Flowchart, Meta-Analysis 2