

Individual Participant Data Meta-Analysis:
Impact of Conduct Problem Severity and Complexity on Parenting Program Effects

Patty Leijten

University of Amsterdam

Stephen Scott

King's College London

Sabine Landau

King's College London

Victoria Harris

Joanna Mann

University of Oxford

Judy Hutchings

Bangor University

Jennifer Beecham

Kent University

Frances Gardner

University of Oxford

Corresponding author: Patty Leijten, University of Amsterdam, PO Box 15780, 1001 NG
Amsterdam, the Netherlands. E: p.leijten@uva.nl; T:+31640236511.

Abstract

Objective: There is concern whether established parenting programs for children's conduct problems meet the needs of families with severe and complex mental health problems. For example, many children with conduct problems show comorbid ADHD or emotional problems, or have parents who are depressed, but families with such complex mental health problems typically seen in real life are often underrepresented in evaluation trials. We tested whether children with more severe conduct problems, and those with more complex mental health problems, benefit less from the Incredible Years parenting program, using individual participant data meta-analysis of randomized trials in Europe. **Method:** In 1,696 families from 13 trials (child age 2–11; 37% girls; 58% low income; 30% ethnic minority; 98% mothers), we used moderator analysis within a multilevel model to test whether initial conduct problem severity, comorbid ADHD or emotional problems and maternal depression diminished intervention effects for children's conduct problems. **Results:** The Incredible Years program reduced children's conduct problems overall (Cohen's $d = -0.35$), but more so in children with more severe conduct problems. There was no evidence that children's comorbid ADHD and emotional problems changed the intervention benefits. Children of mothers with more depressive symptoms benefited more. **Conclusion:** Children with more severe conduct problems derive greater, rather than lesser, benefits from a high quality group parenting program, and comorbid ADHD and emotional problems do not reduce effects; depressed parents, rather than being linked to less child change, were associated with greater reductions in children's conduct problems.

Keywords: conduct problems; parenting program; comorbidity; parental depression; individual participant data meta-analysis.

Individual Participant Data Meta-Analysis:

Impact of Conduct Problem Severity and Complexity on Parenting Program Effects

A major concern about research on the effectiveness of interventions for children's mental health problems is that most families who are recruited to trials have less severe and complex problems than families in clinical practice [1,2]. This is problematic because it hinders generalizability of research findings to the broader population of families seeking support for children's mental health problems. A rigorous test of whether mental health problem severity and complexity attenuates intervention effects requires sufficient variation in problem severity and complexity across families, and sufficient statistical power to detect differential intervention effects by problem severity and complexity. Individual trials rarely meet these criteria: their samples tend to be too homogeneous and too small [3,4]. We therefore pooled individual participant data from a near complete set of 13 European trials on the Incredible Years parenting program for children's conduct problems, whose participants ranged from severe clinically referred cases with multiple comorbidities to early intervention cases with lower levels of problems. We capitalized on the variation and statistical power provided by this combined sample of almost 1,700 families to test whether children with more severe and complex mental health problems benefit less from the program.

Children's conduct problems in early and middle childhood cover a wide range of behaviors, including defiance, temper tantrums, aggression, and destructiveness [5]. These problems cause significant burden and, if left untreated, come with significant societal costs [6,7]. Parenting programs are the recommended strategy to reduce conduct problems [8]. Most established programs are based on social learning theory perspectives and guide parents in breaking coercive parent-child interaction patterns where parents and children unwittingly reinforce aversive behavior in each other in a way that creates cycles of interactions that become increasingly difficult to manage [9]. The Incredible Years parenting program [10] is

one of these established programs. It has a solid evidence-base for its effectiveness in reducing children's conduct problems [11], and is recommended by clearinghouses such as Blueprints (United States) and NICE Guidelines (United Kingdom). The content of the Incredible Years program (i.e., the parenting techniques taught) is similar to that of most other established parenting programs [12], but the delivery method, in particular its strong emphasis on a collaborative group approach and intensive therapist training and supervision, differs from most other parenting programs. The Incredible Years program has been implemented as care as usual in several countries, and yields robust effects across countries [13].

More severe conduct problems are associated with poorer outcomes across several domains of functioning, from juvenile delinquency and criminal violence to leaving school without qualifications and dependence on state benefits [14]. They are therefore a serious mental health problem which is important to treat. However, it is often believed that they will be harder to reduce than milder conduct problems, especially in group programs, since they are often associated with families facing a range of challenges that prevent them from engaging well with interventions [15]. We set out to test this assumption in this study.

Similar concerns exist about comorbid mental health problems. Comorbid mental health problems (e.g., ADHD and emotional problems) are common in children with conduct problem [16,17] and predict more serious deviation from healthy development [18]. Further, mental health problems not only cluster within individuals, but also within families. Many parents of children with conduct problems suffer from depression [19]. In the present study, we tested whether children's comorbid ADHD and emotional problems, and parental depression, attenuate parenting program effects on children's conduct problems.

On the one hand, children with comorbid mental health problems may benefit as much as other children from parenting programs for children's conduct problems. Many

aspects of parenting that are targeted in these programs, such as improving the parent-child relationship and clear and consistent household rules, are important for general child development [20]. On the other hand, children with comorbid mental health problems may benefit less from parenting programs because the program does not target the factors that underlie or maintain their mental health problems. The logic model underlying most established parenting programs is that children's conduct problems are maintained by coercive parenting child interactions. In line with this, families who show more signs of coercion are twice as likely to benefit from parenting programs than other families [21]. When problems are more complex, in terms of comorbid mental health problems, there is a greater likelihood that coercive interactions may not be the primary factor underlying children's conduct problems. For example, conduct problems in some children may be secondary to ADHD [22], or have their origin in internalizing, emotional problems [23]. These children may benefit less from parenting programs, because factors underlying emotional problems and ADHD are not explicitly addressed in parenting programs that focus on breaking cycles of coercive parent-child interactions.

Co-occurring parental mental health problems may attenuate parenting program effects in a different way. Depression can be debilitating, making it more difficult for parents to actively engage in the program, and to work on the skills at home [24]. Because parenting programs such as the Incredible Years rely on the parent to initiate and maintain change in parent-child interactions (i.e., the child itself is not part of the intervention), factors that hinder the parents ability to do this could compromise parenting program effects.

Empirically, there is little evidence that intervention effects are smaller in families with more complex mental health problems. Most individual trials [25,26] and systematic reviews of trials [27,28] find no differential effects for children with or without comorbid mental health problems, and some suggest that children with comorbid mental health

problems benefit more [29–31]. Some recent meta-analyses and larger individual trials suggest that comorbid mental health problems attenuate intervention effectiveness, but this evidence mainly comes from children with anxiety and comorbid mental health problems [32,33]. Findings regarding co-occurring parental mental health problems are similarly inconsistent: Most trials and reviews indicate no differential effects [34,35], and some suggest that families with more depressed parents benefit more [36,37]. Findings regarding problem severity are more consistent. The literature predominantly suggests that children with more severe problems, and thus a larger scope for improvement, benefit more [38,39], although most trials and traditional meta-analyses reporting this finding are underpowered and the literature may suffer from publication bias [3,4].

It can be difficult to detect whether mental health problem severity and complexity attenuate intervention effectiveness. First, problem severity and complexity often go hand in hand: comorbid mental health problems are more prevalent in children with more severe conduct problems [18]. Traditional aggregate level meta-analyses suggest that children with more severe conduct problems tend to benefit more from interventions [11,39], and thus the impact of co-occurrence of mental health problems could be masked by the impact of problem severity. Second, because interaction effects (i.e., participant characteristic \times intervention effects) tend to be smaller than the main effect of interventions, we can only detect such interaction with larger sample sizes. Most individual trials are well-powered to test main effects of interventions, but not interaction effects [3,4]. Findings from moderator analysis via meta-regression tend to be even more difficult to interpret since such an approach is based on aggregate trial level information (e.g., % of children in a sample with comorbid emotional problems), and is therefore only able to detect trial-level effects which might be prone to confounding by other trial level factors, such as program implementation quality. In

addition, the power of such investigations might be low due to limited variability in the moderator summary across trials.

We aimed to overcome these challenges by using pooled individual participant data from multiple randomized trials, to allow for sufficient variation in children's levels of problem severity and complexity, and for sufficient statistical power, to enable us to adequately control for putative confounding moderators. Specifically, we tested the possible impact on parenting program effects of children's conduct problem severity and three aspects of problem complexity: children's comorbid ADHD problems, children's comorbid emotional problems, and parental depression.

Methods

Protocol, Registration, and Reporting

We published our study protocol online (<http://www.spi.ox.ac.uk/parentingIPD>). Procedures were approved by the Departmental Research Ethics Committee of (the Department of Social Policy and Intervention, University of Oxford). We followed PRISMA-IPD guidelines for reporting individual participant data meta-analyses [40].

Identifying and Selecting Trials

We aimed to include all randomized trials of the effects of the Incredible Years parenting intervention in Europe. We chose to focus on European trials because these are all conducted independent of the program developer, and allow for sufficient homogeneity in the usual services that children receive across trials—some family programs developed outside of Europe do not work well in Europe, potentially because of differences in usual services [41].

We identified trials through systematic searches CINAHL, Embase, Global Health, MEDLINE, and PsycINFO, the Incredible Years website overview of trials, the European Incredible Years mentors' network, and asking experts. Eligibility was assessed by the senior author and double checked by four additional authors. More details on study identification are

published elsewhere [42], as are analyses of how family socioeconomic status and children's age impact program effects [43,44].

Included Trials and Participants

We identified 13 eligible trials. This is a near complete set: of the 15 European trials, one did not include parent-reported measures of children's conduct problems because of the children's young age (i.e., toddlerhood [45]) and one trial had not retained the data [46]).

The combined sample included 1,696 families (1,046 intervention; 650 control condition—some trials used a 2:1 allocation ratio). Children ranged from 2 to 10 years of age ($M = 5.26$; $SD = 1.49$; 37% girls). Families were diverse in terms of socioeconomic status (58% low income; 63% low educational level; 35% unemployed; 35% single parent; 12% teen parent) and ethnic background (31% minority). Most trials included data from only one parent (98% mothers).

Rates of conduct problem severity and complexity varied widely across trials, in part because we deliberately combined prevention and treatment trials—to increase variation in problem severity and complexity, and in part because trials were conducted in very different settings (e.g., schools in inner city London, psychiatric clinics in Norway and Sweden, and community settings in the Netherlands; Table 1). As a consequence, scores for children's conduct problems, ADHD, emotional problems, and parental depressive symptoms covered almost the full possible range (Table 2). Of the children, 38% showed clinical levels of ADHD symptoms (Strengths and Difficulties Questionnaire hyperactivity/inattention score >7) and 28% showed clinical levels of emotional problems (Strengths and Difficulties Questionnaire emotional problems score >5) [47]; of the parents (98% mothers), 20% showed clinical levels of depression (Beck Depression Inventory score >30) [48].

Parents in the intervention conditions attended on average 63% of the sessions (range 0–100%; $SD = 35\%$). Correlations between attendance and maternal and child mental health

were weak, although some yielded significance because of the large sample size (maternal depression: $r = -.082$, $n = 684$, $p = .031$; conduct problems: $r = .087$, $n = 831$, $p = .012$; ADHD: $r = .045$, $n = 757$, $p = .219$; emotional problems $r = -.034$, $n = 662$, $p = .386$).

Measures

Conduct problems. Most trials measured children's conduct problems in using the Intensity Scale of the Eyberg Child Behavior Inventory (ECBI [49]); two trials used the Parental Account of Children's Symptoms (PACS [50]). PACS scores were converted into ECBI scores using norm deviation scores (see [42], for more details on this procedure). ECBI and PACS scores correlated $r = .71$ in our sample, based on data from four trials that included both measures. Internal consistency ranged from $\alpha = .79$ to $.95$. All 13 trials contributed data ($n = 1,622$).

ADHD symptoms. Most trials measured children's ADHD symptoms using the Hyperactivity/Inattention scale of the Strengths and Difficulties Questionnaire (SDQ [51]); two trials used the Child Behavior Checklist (CBCL [52]); and one trial used the PACS. CBCL and PACS scores were converted to SDQ scores using norm deviation scores. Eleven trials contributed data ($n = 1,532$).

Emotional problems. Most trials measured children's emotional problems using the Emotional Symptoms scale of the SDQ; two trials used the CBCL; and one trial used the PACS. CBCL and PACS scores were converted to SDQ scores using norm deviation scores. Ten trials contributed data ($n = 1,340$).

Maternal depression. Most trials measured parental depressive symptoms using the Beck Depression Inventory ([53]); one trial used the Brief Symptom Inventory—depression subscale (BSI [54]); and one trial used the General Health Questionnaire (GHQ [55]). BSI and GHQ scores were converted to BDI scores using norm deviation scores. Internal

consistency ranged from $\alpha = .87$ to $.93$. Eleven trials contributed data ($n = 1,395$). Data from 98% of the families come from mothers.

Risk of Bias

We assessed risk of bias in the trials as high, low, or unclear using the Cochrane risk of bias tool. Specifically, we assessed potential risk of bias concerning random sequence generation, allocation concealment, blinding of assessors, addressing incomplete data, selective outcome reporting, and other sources of bias. Risk of bias was low on most indicators for most trials, with the exception of blinding of assessors. As is typical in parenting program evaluation studies, parents were aware they were participating in a parenting program, and were the main informants of program effects.

Analytic Strategy

All analyses were pre-specified; we published our analysis plan online ahead of conducting the analyses (<http://www.spi.ox.ac.uk/parentingIPD>).

Correlations between moderators. We used pairwise Pearson's correlation coefficients calculated on the basis of the available data to measure how strongly each of the putative moderators co-vary, and to understand the co-occurrence of complex problems.

Modelling approach. We used multilevel modeling (random effect modeling) to capture the hierarchical structure of the data, with families (Level 1) nested within Incredible Years therapy groups (Level 2) within the intervention condition, and therapy groups nested within trials (Level 3). We took trial design features (e.g., cluster and stratified randomization) into account by including trial-specific fixed effects. In addition, to allow for further intervention effect heterogeneity (e.g., due to unmeasured trial characteristics) we included a trial-varying random coefficient of condition in the model. We fitted models by maximum likelihood, which is valid under a missing at random (MAR) assumption about the missing data. All analyses were carried out in Stata version 14 and were based on intention-

to-treat principles: participants were analyzed in the conditions to which they were randomized, irrespective of whether or how much they participated in Incredible Years. Because the functional form of the relation between the outcome and moderators is not known theoretically, we assessed this empirically by testing for a non-linear relationship between the outcome and each moderator (these were all continuous variables), by adding a quadratic term and a condition \times quadratic term to the model. If the additional terms significantly improved the fit (at liberal $\alpha = .10$ level), we added them to the model; otherwise we maintained the more parsimonious linear relationship.

First, we modelled each putative moderator separately to determine unadjusted moderator effects. We used children's conduct problems post-intervention as the dependent variable and included fixed effects for condition (i.e., Incredible Years or control), trial level moderator summaries (between-trial variables, e.g., mean ADHD score in a trial), participant level deviations from trial summaries (within-trial variables, e.g., individual participant ADHD score), and respective interaction terms. Including interaction terms at both the trial and participant level allowed us to assess empirically whether these two moderating effects differed. In other words, it allowed us to see whether moderator analysis using individual participant data meta-analysis yielded different findings than traditional moderator analysis using a trial level meta-analytic approach. If this difference was significant at a liberal $\alpha = .10$, we interpreted both effects; if not, we interpreted the more powerful model with a single interaction term. Moderation effect sizes were expressed as the estimated intervention effect on the ECBI per one unit standard deviation of the putative moderator. For example, a moderation index of -2 indicates that with every unit standard deviation increase on the moderator variable, the effect of the parenting program enlarges by an average two points on the ECBI.

Second, we investigated adjusted moderator effects by expanding models to include further moderator \times condition interaction terms for any variable that correlated significantly with the target moderator and that was found to be a moderator in the unadjusted analysis. If adding these interaction terms reduces the target interaction effect, it means that the adjustment variable could account at least partly for the observed unadjusted moderation effect.

Missing data. Missing data were mainly due to the fact that not all trials included all putative moderator variables (Table 2). We used Multiple Imputation to minimize missing data biases by including observed predictors of missing values in the moderator variable in the imputation step of the Multiple Imputation procedure. We used binary logistic regression to identify baseline demographic variables that predicted missingness of each putative moderator, controlling for trial, condition, child gender and age and baseline ECBI score, because these variables were already included in the imputation step for other reasons. Baseline emotional problems predicted missingness for ADHD symptoms and teenage parenthood predicted missingness for maternal symptoms of depression. These predictors were therefore included in the imputation to ensure that the model was valid under a more realistic ‘missing at random’ assumption regarding the missing data generating process.

Power analysis. Because of the large pooled sample size for each analyses ($n = 1,340$ to $n = 1,622$), statistical power was more than adequate. Power was up to 96% to detect small (Cohen’s $d = .20$) and 80% to detect very small ($d = .15$) condition \times severity or complexity interaction effects at $\alpha = 0.05$.

Results

Descriptive Analyses

ADHD symptoms and emotional problems correlated strongly and moderately respectively, with conduct problem severity ($r = .41, n = 1,495$ and $r = .26, n = 1,322$).

ADHD symptoms and emotional problems also correlated moderately with each other ($r = .26$, $n = 1,318$). Maternal depression correlated positively with children's mental health problems (conduct problems $r = .27$, $n = 1,326$; ADHD $r = .18$, $n = 1,241$; emotional problems $r = .28$, $n = 1,072$). All correlations were highly significant ($p < .001$). Thus, problem severity and problem complexity variables were positively correlated and could potentially explain moderating effects for any one of these variables in an unadjusted analysis.

The average effect of the parenting program on children's conduct problems was 13.5 points on ECBI (95% CI 10.9 to 16.1), indicating a standardized effect size of $\beta = 0.35$, 95% CI 0.51 to 0.19). There was much heterogeneity underlying this average effect: 44% of children in the intervention condition showed reliable improvement [56] in conduct problems (versus 24% of children in the control condition) and 6% of children in the intervention condition showed reliable worsening (versus 9% of children in the control condition). In other words, children varied substantially from each other in how much they benefited from the parenting program, highlighting the need for moderator analysis.

Primary Analyses

Conduct problem severity as moderator. There was evidence that any program effect moderation by children's baseline levels of conduct problems varied between the trial and individual participant level ($p = .004$). There was no evidence that the functional relation between baseline levels of conduct problems and post-intervention conduct problems was not linear ($p = .09$). At the trial level, there was a large significant moderation effect (moderator effect size -18.3 ECBI points, 95% CI -24.6 to -12.0 ; $p = .001$). At the individual participant level, there was a more modest significant moderation effect (moderator effect size -4.3 ECBI points, 95% CI -7.9 to -0.7 points; $p = .02$; Figure 1). Thus, both at the trial and

individual participant level, children with higher levels of conduct problems at baseline benefited more.

ADHD symptoms as moderator. There was no evidence that any program effect moderation by children's ADHD symptoms varied between the trial and individual participant level ($p = .58$), but there was a suggestion that the functional relation was not linear ($p = .02$). The moderation effect was, however, not significant ($p = .07$). Therefore, there is insufficient evidence that parenting program effects on children's conduct problems were moderated by children's ADHD symptoms. Children with higher levels of comorbid ADHD symptoms did not benefit significantly less, or more, from the parenting program in terms of reduced conduct problems.

Emotional problems as moderator. There was no evidence that any program effect moderation by children's emotional problems varied between the trial and individual participant level ($p = .28$), or that the functional relation was not linear ($p = .38$). The single moderation effect was not significant (moderator effect size -2.3 ECBI points, 95% CI -6.7 to 0.9 points; $p = .13$). Therefore, there is no evidence to suggest that parenting program effects on children's conduct problems were moderated by children's emotional problems. Children with higher levels of comorbid emotional problems did not benefit less, or more, from the parenting program in terms of reduced conduct problems.

Maternal depressive symptoms as moderator. There was no evidence that any program effect moderation by maternal depressive symptoms varied between the trial and individual participant level ($p = .30$), or that the functional relation was not linear ($p = .31$). The single moderation effect was significant (moderator effect size -4.8 ECBI points, 95% CI -8.0 to -0.9 points, $p = .01$; Figure 2). Children with mothers with higher levels of depressive symptoms benefited more from the parenting program in terms of reduced conduct problems.

Moderator Effects after Adjusting for Confounding

Because the two significant moderator variables (i.e., severity of conduct problems and maternal depressive symptoms) were correlated ($r = 0.27$, $p < .001$, $n = 1326$), we assessed whether the moderator effect of maternal depressive symptoms could be accounted for by conduct problem severity by including interaction effects of both variables in one model. The adjusted moderator effects for maternal depressive symptoms reduced to a marginally significant trend effect (-3.40 ECBI points; $p = .07$), in the same direction as the unadjusted moderator effects (i.e., there was a trend that children with more depressed mothers benefit more). In other words, children with more depressed mothers benefited more from the parenting program, partly accounted for by the on average more severe conduct problems of these children at baseline that increased program effects. Conduct problem severity was no longer a significant moderator after adjusting for moderation by maternal depressive symptoms (-1.74 ECBI points; $p = .36$). In other words, children with more severe conduct problems benefited more from the parenting program, accounted for by the on average higher maternal depression rates in these families that increased program effects.

Discussion

We examined whether children's conduct problems severity and comorbid ADHD and emotional problems and maternal depressive symptoms impact the effects of a parenting program on children's conduct problems, using data from an integrated sample of almost 1700 families participating in randomized trials of the Incredible Years program in Europe. Our findings suggest that children with more comorbid ADHD symptoms or emotional problems benefit as much as children with less comorbid problems. Children with more severe conduct problems and children with more depressed mothers benefited more.

That children with more severe conduct problems benefit more from interventions to reduce conduct problems has previously been suggested by some individual trials and some

trial aggregate level meta-analyses [8,11], although some individual trials suggest opposite patterns [57]. However, never before has the much greater power of individual level participant analysis been harnessed to address this issue, so the present findings add considerable weight to the proposition. In individual participant data meta-analysis, variation within trials is aggregated across all cases, rather than differences in characteristics of all trials being the only way of calculating effects [3,4]. The finding of more severe cases doing better in any observational studies of mental health interventions is sometimes dismissed as regression to the mean, but this cannot explain the findings of this study: regression to the mean would apply equally to both the intervention and the control group. We saw, however, that above and beyond a stronger reduction in conduct problems in all children with more severe conduct problems (i.e., in both trial arms), there was a differentially greater effect of problem severity for children in the intervention group. This effect disappeared in the adjusted model, where we corrected the impact of baseline levels of conduct problems for the impact of baseline levels of maternal depression. That children with more severe conduct problems benefited more for the parenting program was thus partly accounted for by the association with on average higher levels of maternal depression in these families that impacted program effects.

Our finding that children with more severe conduct problems benefit more was significantly stronger at the trial level (moderator effect size -18.3) than at the individual family level (moderator effect size -4.3). Individual participant data meta-analysis provides a more precise estimation of moderator effects—it takes both between- and within trial variance into account and is much better powered. That traditional trial level meta-analysis may overestimate moderator effects, at least in our data, suggests moderator findings of traditional trial level meta-analyses should be interpreted with caution, and replication in individual participant data meta-analysis is advised.

Our findings that children's comorbid ADHD or emotional problems do not impact program benefits may be surprising. Although some systematic reviews suggest the same [27,28], most of these findings are based on individual trials that are severely underpowered for rigorous moderation or subgroup analyses. Of the 1969 children in our sample, more than 644 children had at least one type of comorbid mental health problem. With such a large sample size, we were well-equipped to identify any divergent response patterns in these children. Our findings, however, suggest that, at least in the case of the Incredible Years parenting program, comorbid mental health problems do not seem to stand in the way of effective intervention for children's conduct problems. Besides, there is evidence that parenting programs directly improve ADHD as well [42].

Children of mothers who suffer more from depressive symptoms showed a stronger reduction in conduct problem in reaction to the parenting program. This effect was robust in the face of adjusting for the moderator effect of children's problem severity. Although we were unable to test the mechanisms underlying this effect, there may be several possible explanations for this finding. First, depression can compromise mothers' abilities to be sensitive and consistent towards their children [24] and may therefore come with a larger scope to improve parenting practices. If parenting practices indeed improve more in mothers with more depressive symptoms, this could explain why their children's behavior improved more. Second, participation in a parenting program that uses a collaborative approach to empower parents and a group setting to reduce social isolation, could perhaps relieve depression and loneliness in mothers who suffer from depressive symptoms and provide them with peer support. Thirdly, setting realistic short-term goals in a parenting program, which then produce immediate positive effects on child behavior, can help to lift a parent's mood, via reinforcement mechanisms similar to those operating in behavioral activation for depression [58]. Such experiences would initiate a positive feedback loop encouraging

engagement by depressed mothers in the program. Reduction in parental depression in turn benefits children's mental health [59].

Strengths and Limitations

Strengths of our study include the uniquely large combined sample with sufficient variation in problem severity and complexity across families, allowing us to rigorously test moderation effects with exceptional statistical power. We also tested for potentially confounding moderator effects, vital for understanding the impact of problem complexity because problem severity and complexity often go hand in hand.

Limitations of our study include our use of parent-reported outcomes—parents in randomized trials of parenting programs are not blind to condition. We note though that our goal was to estimate patterns of differential effectiveness (i.e., relative levels of effectiveness), rather than to estimate the magnitude of effectiveness (i.e., absolute levels of effectiveness). In addition, we focused on immediate parenting program effects only, because most trials used a waitlist control condition where families in the control condition received the intervention immediately after post-intervention assessment. Findings from a recent meta-analysis [21] are reassuring in showing that parenting program effects are on average stable in the months and years after the program, but potentially differential longer-term effects by problem severity and complexity are yet to be explored. While ADHD and emotional problems may be the most prevalent co-morbid mental health problems in children with conduct problems, there may be other relevant co-morbid problems we were unable to include in this study (e.g., autism or intellectual disabilities). Lastly, these findings may not generalize to other parenting programs, some of which use different delivery methods than the collaborative group process used in the Incredible Years, and some of which that have less intensive therapist training and implementation quality procedures.

Clinical Implications

Our findings are potentially reassuring for clinical practitioners who may feel that they struggle to achieve change in children with more complex mental health problems. Important to note here, however, is that although comorbid mental health problems do not seem to stand in the way of reducing children's conduct problems, this does not mean that the comorbid problems themselves also reduce. Our pooled data show that ADHD symptoms benefit from this parenting intervention, but emotional problems, albeit in most trials measured by the brief Strengths and Difficulties Questionnaire, appear not to [42]. Recent developments in transdiagnostic conceptualizations [60] and interventions [61] may be helpful in cases where the goal is to reduce multiple types of mental health problems at the same time.

Suggestions for Future Research

While initial problem severity and maternal depression explain to some extent why some children benefit more than others from parenting programs, much heterogeneity in program benefits remains unexplained. One reason for this could be the perhaps oversimplified approach of testing individual child or parent characteristics as putative moderators of program effects. Family characteristics interact in predicting parenting program effects and person-centered approaches that allow family characteristics to cluster in predicting intervention benefits can further advance our understanding of the children that benefit less or more [38,62].

Related to this, the impact of children's comorbid problems on parenting program effects may depend on the nature of children's comorbid problems, such as whether children's conduct problems are at the basis of some of the other problems, or whether they are the consequence of other problems [63]. We therefore encourage the field to explore ways to take children's developmental history into account when exploring the role of comorbid mental health problems in intervention effectiveness.

Conclusion

Our findings suggest that it is not more difficult to reduce conduct problems in children whose mental health problems are more severe or complex. If anything, children with more severe conduct problems, and those with more depressed mothers, seem to benefit more, and children with comorbid ADHD or emotional problems fare just as well. Next steps for advancing our understanding of how comorbid mental health problems impact treatment effectiveness include studying how processes underlying comorbid mental health problems impact program effects.

References

1. Kazdin AE, Wassell G. Therapeutic changes in children, parents, and families resulting from treatment of children with conduct problems. *J Am Acad Child Adolesc Psychiatry*. 2000;39:414–420.
2. Weisz J, Ng M, Lau N. Psychological interventions: Overview and critical issues for the field. In A Thapar, D Pine, J Leckman, S Scott, M Snowling, E Taylor, eds. *Rutter's Child and Adolescent's Psychiatry* (6th ed.). Oxford and New York: Wiley Blackwell. 2015.
3. Fisher DJ, Carpenter JR, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;356:573–578.
4. Thompson SG, Higgins JPT. Can meta-analysis help target interventions at individuals most likely to benefit? *Lancet*. 2005;365:341–346.
5. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association. 2013.
6. Romeo R, Knapp M, Scott S. Economic cost of severe antisocial behavior in children—and who pays it. *Br J Psychiatry*. 2006;188:547–553.

7. Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: Follow up study of antisocial children into adulthood. *BMJ*. 2001;323:191–195.
8. Weisz JR, Kazdin AE. Evidence-based psychotherapies for children and adolescents. New York: Guildford Press; 2018.
9. Patterson, GR. Coercive family process. Eugene: Castalia; 1982.
10. Webster-Stratton C, Reid M. The Incredible Years parents, teachers, and children training series: a multifacted treatment approach for children with conduct disorders. In: Weisz J, Kazdin A, eds. Evidence-Based Psychotherapies for Children and Adolescents. New York, NY: Guilford Press; 2010:194–210.
11. Menting ATA, Orobio de Castro B, Matthys W. Effectiveness of the Incredible Years parent training to modify disruptive and prosocial child behavior: A meta-analytic review. *Clin Psychol Rev*. 2013;33:901–913.
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;19:236–258.
13. Leijten P, Melendez-Torres GJ, Knerr W, Gardner F (2016). Transported versus homegrown parenting interventions for reducing disruptive child behavior: A multilevel meta-regression study. *J Am Ac Child Adolesc Psychiatry*. 2016;55:610–617.
14. Fergusson D, Horwood L, Ridder E. Show me a child at seven: Consequences of conduct problems in childhood for psychosocial functioning in adulthood. *J Child Psychol Psychiatry*. 2005;46:837–849.
15. Nock MK, Ferriter C. Parent management of attendance and adherence in child and adolescent therapy: A conceptual and empirical review. *Clin Child Fam Psychol Rev*. 2005;8:149–166

16. Green JG, McLaughlin KA, Berglund PA, Gruber MJ, Sampson NA, Zaslavsky AM et al. Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: associations with first onset of DSM-IV disorders. *Arch Gen Psychiatry*. 2010;67:113–123.
17. Polier GG, Vloet TD, Herpertz-Dahlmann B, Laurens KR, Hodgins S. Comorbidity of conduct disorder symptoms and internalising problems in children: investigating a community and a clinical sample. *Eur Child Adolesc Psychiatry*. 2012;21:31–38.
18. Rutter M, Kim-Cohen J, Maughan B. Continuities and discontinuities in psychopathology between childhood and adult life. *J Child Psychol Psychiatry*. 2006;47:276–295.
19. Gross H, Shaw DS, Burwell RA, Nagin DS. Transactional processes in child disruptive behavior and maternal depression: A longitudinal study from early childhood to adolescence. *Dev Psychopathol*. 2009;21:139–156.
20. Bornstein MH. Children's parents. In: Bornstein MH, Leventhal T, editors. *Ecological settings and processes in developmental systems*. Hoboken, NJ: Wiley; 2015.55–132.
21. Van Aar J, Leijten P, Orobio de Castro B, Overbeek G. Sustained, fade-out or sleeper effects? A multilevel meta-analysis of parenting interventions for disruptive child behavior. *Clin Psychol Rev*. 2017;51:153–163.
22. Danforth JS, Connor, DF, Doerfler LA. The development of comorbid conduct problems in children with ADHD: an example of an integrative developmental psychopathology perspective. *J Atten Disord*. 2016;20:214–229.
23. Vieno A, Kiesner J, Pastore M, Santinello M. Antisocial behavior and depressive symptoms: Longitudinal and concurrent relations. *Adolescence*. 2008;43:649–660.
24. Lovejoy MC, Graczyk PA, O'Hare E, Neuman G. Maternal depression and parenting behavior: A meta-analytic review. *Clin Psychol Rev*. 2000;20:561–592.

25. Børnebekk G, Kjøbli J, Ogden T. Children with conduct problems and co-occurring ADHD: Behavioral improvements following parent management training. *Child Fam Behav Ther.* 2015;37:1–19.
26. Farmer CA, Brown NV, Gadow KD, Arnold LE, Kolko DG, Findling RL et al. Comorbid symptomatology moderates response to risperidone, stimulant, and parent training in children with severe aggression, disruptive behavior disorder, and attention-deficit/hyperactivity disorder. *J Child Adolesc Psychopharmacol.* 2015;25:213–224.
27. Ollendick TH, Jarrett MA, Grills-Tauechel AE, Hovey LD, Wolff JC. Comorbidity as a predictor and moderator of treatment outcome in youth with anxiety, affective, attention deficit/hyperactivity disorder, and oppositional/conduct disorders. *Clin Psychol Rev.* 2008;28:1447–1471.
28. Riosa PB, McArthur BA, Preyde M. Effectiveness of psychosocial intervention for children and adolescents with comorbid problems: A systematic review. *Child Adolesc Ment Health.* 2011;16:177–185.
29. Jarrett M, Siddiqui S, Lochman J, Qu L. Internalizing problems as a predictor of change in externalizing problems in at-risk youth. *J Clin Child Adolesc Psychol.* 2014;43:27–35.
30. Kazdin AE, Whitley MK. Comorbidity, case complexity, and effects of evidence-based treatment for children referred for disruptive behavior. *J Consult Clin Psychol.* 2006;74:455–467.
31. Leijten P, Raaijmakers MAJ, Wijngaards L, Matthys W, Menting ATA, Hemink-van Putten M, Orobio de Castro B. Understanding who benefits from parenting interventions for children's conduct problems: An integrative data analysis. *Prev Sci.* 2018;19:579–588.

32. Rapee RM, Lyneham HJ, Hudson JL, Kangas M, Wuthrich VM, Schniering CA. Effect of comorbidity on treatment of anxious children and adolescents: results from a large, combined sample. *J Am Acad Child Adolesc Psychiatry*. 2013;52:47–56.
33. Walczak M, Ollendick T, Ryan S, Esbjørn BH. Does comorbidity predict poorer treatment outcome in pediatric anxiety disorders? An updated 10-year review. *Clin Psychol Rev*. 2018;60:45–61.
34. Baydar N, Reid MJ, Webster-Stratton C. The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for Head Start mothers. *Child Dev*. 2003;74:1433–1453.
35. Reyno SM, McGrath PJ. Predictors of parent training efficacy for child externalizing behavior problems—a meta-analytic review. *J Child Psychol Psychiatry*. 2006;47:99–111.
36. Gardner F, Hutchings J, Bywater T, Whitaker C. Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *J Clin Child Adolesc Psychol*. 2010;39:568–580.
37. Shaw DS, Dishion TJ, Supplee L, Gardner F, Arnds K. Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-year effects of the family check-up in early childhood. *J Consult Clin Psychol*. 2006;74:1–9.
38. Van Aar J, Leijten, P, Orobio de Castro B, Weeland J, Matthys W, Chhangur R, Overbeek G. 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. <https://doi.org/10.1016/j.jaac.2019.02.004>
39. Leijten P, Gardner F, Melendez-Torres GJ, Van Aar J, Hutchings J, Schulz S et al. Meta-analyses: Key parenting program components for disruptive child behavior. *J Am Acad Child Adolesc Psychiatry*. 2019;58:180–190.

40. Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G et al. Preferred reporting items for a systematic review and meta-analysis of individual participant data: The PRISMA-IPD statement. *JAMA*. 2015;313:1657–1665.
41. Robling M, Bekkers MJ, Bell K, Butler CC, Cannings-John R, Channon S et al. Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (Building Blocks): a pragmatic randomised controlled trial. *The Lancet*. 2016;387:146–155.
42. Leijten P, Gardner F, Landau S, Harris V, Mann J, Hutchings J. et al. 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*. 2008;59:99–109.
43. Gardner F, Leijten P, Harris V et al.. Equity effects of parenting interventions for child conduct problems. *The Lancet Psychiatry*. 2019;Published online.
44. Gardner F, Leijten P, Melendez-Torres GJ, Scott S, Beecham J, Mann J et al. Earlier the better? Individual participant data and traditional meta-analysis of age effects of parenting intervention. *Child Dev*. 2019;90:7–19.
45. Patterson J, Barlow J, Mockford C, Klimes I, Pyper C, Stewart-Brown S. Improving mental health through parenting programmes: block randomised controlled trial. *Arch Dis Child*. 2002;87:472–477.
46. Hutchings J, Griffith N, Bywater T, Williams ME. Evaluating the Incredible Years Toddler Parenting Programme with parents of toddlers in disadvantaged (Flying Start) areas of Wales. *Child Care Health Dev*. 2017;43:104–113.
47. National Health Interview Survey [NHIS] (2001). Centers for Disease Control and Prevention National Center for Health Statistics. Available at: <http://www.cdc.gov/nchs/nhis.htm>.

48. Veerman JL, Dowrick C, Ayuso-Mateos JL, Dunn G, Barendregt JJ. Population prevalence of depression and mean Beck Depression Inventory Score. *Br J Psychiatry*. 2009;195:516–519.
49. Eyberg SM, Ross AW. Assessment of child behavior problems: The validation of a new inventory. *J Clin Child Psychol*. 1987;7:113–116.
50. Taylor E, Schachar R, Thorley G, Wieselberg M. Conduct disorder and hyperactivity: I. Separation of hyperactivity and antisocial conduct in British child psychiatric patients. *Brit J Psychiatry*. 1986;149:760–767.
51. Goodman R. The Strengths and Difficulties Questionnaire: A research note. *J Child Psychol Psychiatry*. 1997; 38:581–586.
52. Achenbach TM. Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington, VT: University of Vermont; 1991.
53. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh, J. An inventory for measuring depression. *Arch Gen Psychiatry*. 1961;4:561–571.
54. Derogatis LR, Spencer PM. The Brief Symptom Inventory: Administration, Scoring and Procedures Manual 1. Baltimore, MD: Clinical Psychometrics Research; 1982.
55. Williams P, Goldberg DP. A user's guide to the General Health Questionnaire. Berkshire: NFER, Nelson; 1988.
56. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59:12–19.
57. Kazdin AE, Wassell G. Therapeutic changes in children, parents, and families resulting from treatment of children with conduct problems. *J Am Acad Child Adolesc Psychiatry*. 2000;39:414–420.
58. Cuijpers P, Van Straten A, Warmerdam L. Behavioral activation treatments of depression: A meta-analysis. *Clin Psychol Rev*. 2007;27:318–326.

59. Cuijpers P, Weitz E, Karyotaki E, Garber J, Andersson G. The effects of psychological treatment of maternal depression on children and parental functioning: a meta-analysis. *Eur Child Adolesc Psychiatry*. 2015;24:237–245.
60. Fraire MG, Ollendick TH. Anxiety and oppositional defiant disorder: A transdiagnostic conceptualization. *Clin Psychol Rev*. 2013;33:229–240.
61. Weisz J, Bearman SK, Santucci LC, Jensen-Doss A. Initial test of a principle-guided approach to transdiagnostic psychotherapy with children and adolescents. *J Clin Child Adolesc Psychol*. 2017;46:44–58.
62. Pelham WE, Dishion TJ, Tein JY, Shaw DS, Wilson MN. What doesn't work for whom? Exploring heterogeneity in responsiveness to the family check-up in early childhood using a mixture model approach. *Prev Sci*. 2017;18:911–922.
63. Cunningham NR, Ollendick TH. Comorbidity of anxiety and conduct problems in children: Implications for clinical research and practice. *Clin Child Fam Psychol Rev*. 2010;13:333–347.
64. Axberg U, Broberg AG. Evaluation of “The Incredible Years” in Sweden: The transferability of an American parent-training program to Sweden. *Scand J Psychol*. 2012;53:224–232.
65. Azevedo A, Seabra-Santos MJ, Gaspar MF, Homem TC. The Incredible Years Basic Parent Training for Portuguese preschoolers with AD/HD behaviours: does it make a difference? *Child Youth Care Forum*. 2013;42:403–424.
66. Gardner F, Burton J, Klimes I. Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: outcomes and mechanisms of change. *J Child Psychol Psychiatry*. 2006;47:1123–1132.

67. Hutchings J, Bywater T, Daley D, Gardner F, Whitaker C, Jones K et al. Parenting intervention in Sure Start services for children at risk of developing conduct disorder: Pragmatic randomised controlled trial. *BMJ*. 2007; 334:678–682.
68. Larsson B, Fossum S, Clifford G, Drugli MB, Handegård BH, Mørch WT. Treatment of oppositional defiant and conduct problems in young Norwegian children: results of a randomized controlled trial. *Eur Child Adolesc Psychiatry*. 2009;18:42–52.
69. Leijten P, Raaijmakers MAJ, Orobio de Castro B, Van den Ban E, Matthys W. Effectiveness of the Incredible Years parenting intervention for families with socioeconomically disadvantaged and ethnic minority backgrounds. *J Clin Child Adolesc Psychol*. 2017;46:59–73.
70. McGilloway S, Mhaille GN, Bywater T, Furlong M, Leckey Y, Kelly P et al. A parenting intervention for childhood behavioral problems: A randomised controlled trial in disadvantaged community-based settings. *J Consult Clin Psychol*. 2012;80:116–127.
71. Menting ATA, Orobio de Castro B, Wijngaards-de Meij, L, Matthys W. A trial of parent training for mothers being released from incarceration and their children. *J Clin Child Adolesc Psychol*. 2014;43:381–396.
72. Morpeth L, Blower S, Tobin K, Taylor RS, Bywater T, Edwards RT et al. The effectiveness of the Incredible Years pre-school parenting programme in the United Kingdom: a pragmatic randomised controlled trial. *Child Care Practice*. 2017; 23:141–161.
73. Scott S, Spender Q, Doolan M, Jacobs B, Aspland H. Multicentre controlled trial of parenting groups for childhood antisocial behaviour in clinical practice. *BMJ*. 2001;323:194–198.

74. Scott S, Sylva K, Kallitsoglou A, Ford T. Which type of parenting programme best improves child behaviour and reading? Follow-up of the helping children achieve trial. London: Nuffield Foundation. 2014.
75. Scott S, O'Connor TG, Futh A, Matias C, Price J, Doolan M. Impact of a parenting program in a high-risk, multi-ethnic community: The PALS trial. *J Child Psychol Psychiatry*. 2010;51:1331–1341.
76. Scott S, Sylva K, Doolan M, Price J, Jacobs B, Crook C, Landau S. Randomised controlled trial of parent groups for child antisocial behaviour targeting multiple risk factors: the SPOKES project. *J Child Psychol Psychiatry*. 2010;51:48–57.
77. 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;23:15–28.

Table 1. *Individual Trial Characteristics.*

Trial	Country	Setting	<i>N</i>	Age (<i>M</i>)	% ethnic minority	% low income	Conduct problems <i>M (SD)</i>
Axberg et al. (2012) [64]	Sweden	Outpatient psychiatric clinics	62	3–8 (5.97)	0	41	155.01 (22.15)
Azevedo et al. (2013) [65]	Portugal	University clinics	124	3–6 (4.66)	0	0	127.61 (28.71)
Gardner et al. (2006) [66]	England	Community services	76	2–9 (5.93)	2	64	161.47 (37.21)
Hutchings et al. (2017) [67]	Wales	Community services	153	3–4 (3.84)	1	79	145.06 (26.98)
Larsson et al. (2009) [68]	Norway	Outpatient psychiatric clinics	75	3–8 (6.58)	1	25	158.04 (23.94)
Leijten et al. (2017) [69]	Netherlands	Outpatient psychiatric clinics & schools	156	2–8 (5.59)	65	74	124.24 (32.83)

McGilloway et al. (2012) [70]	Ireland	Community services	149	2–7 (4.84)	6	47	158.54 (30.52)
Menting et al. (2014) [71]	Netherlands	Community services for formerly incarcerated mothers	99	1–11 (6.30)	78	93	109.66 (31.21)
Morpeth et al. (2017) [72]	England	Community services	161	2–4 (3.68)	52	63	143.08 (36.53)
Scott et al. (2001) [73]	England	Outpatient psychiatric clinics	141	2–10 (5.67)	15	58	162.52 (29.75)
Scott et al. (2014) [74]	England	Schools	214	3–7 (6.07)	19	80	136.13 (31.70)
Scott, O'Connor et al. (2010) [75]	England	Schools	174	4–6 (5.50)	75	44	104.78 (33.42)
Scott, Sylva et al. (2010) [76]	England	Schools	112	4–6 (5.21)	40	44	128.42 (42.96)

Table 2. *Aggregate Trial Baseline Levels of Problem Severity and Complexity.*

	Incredible Years					Control			
	<i>k</i>	<i>n</i>	Range (possible)	<i>M</i>	<i>SD</i>	Per cent in clinical range	<i>M</i>	<i>SD</i>	Per cent in clinical range
<i>Problem severity</i>									
Conduct problems	13	1622	44–252 (36–252)	139.4	37.0	53% ^a	135.5	37.0	49% ^a
<i>Problem complexity</i>									
Child ADHD symptoms	11	1532	0–10 (0–10)	5.9	2.7	36% ^b	5.8	2.7	40% ^b
Child emotional problems	10	1340	0–10 (0–10)	3.4	2.7	30% ^b	3.2	2.4	26% ^b
Parental depression	11	1395	0–59 (0–63)	12.2	10.9	22% ^c	10.1	9.7	19% ^c

Note. *k* = number of trials contributing data; *n* = number of participants. ^a >80th percentile [77]; ^b “abnormal” [48]; ^c “clinical rating” [49].

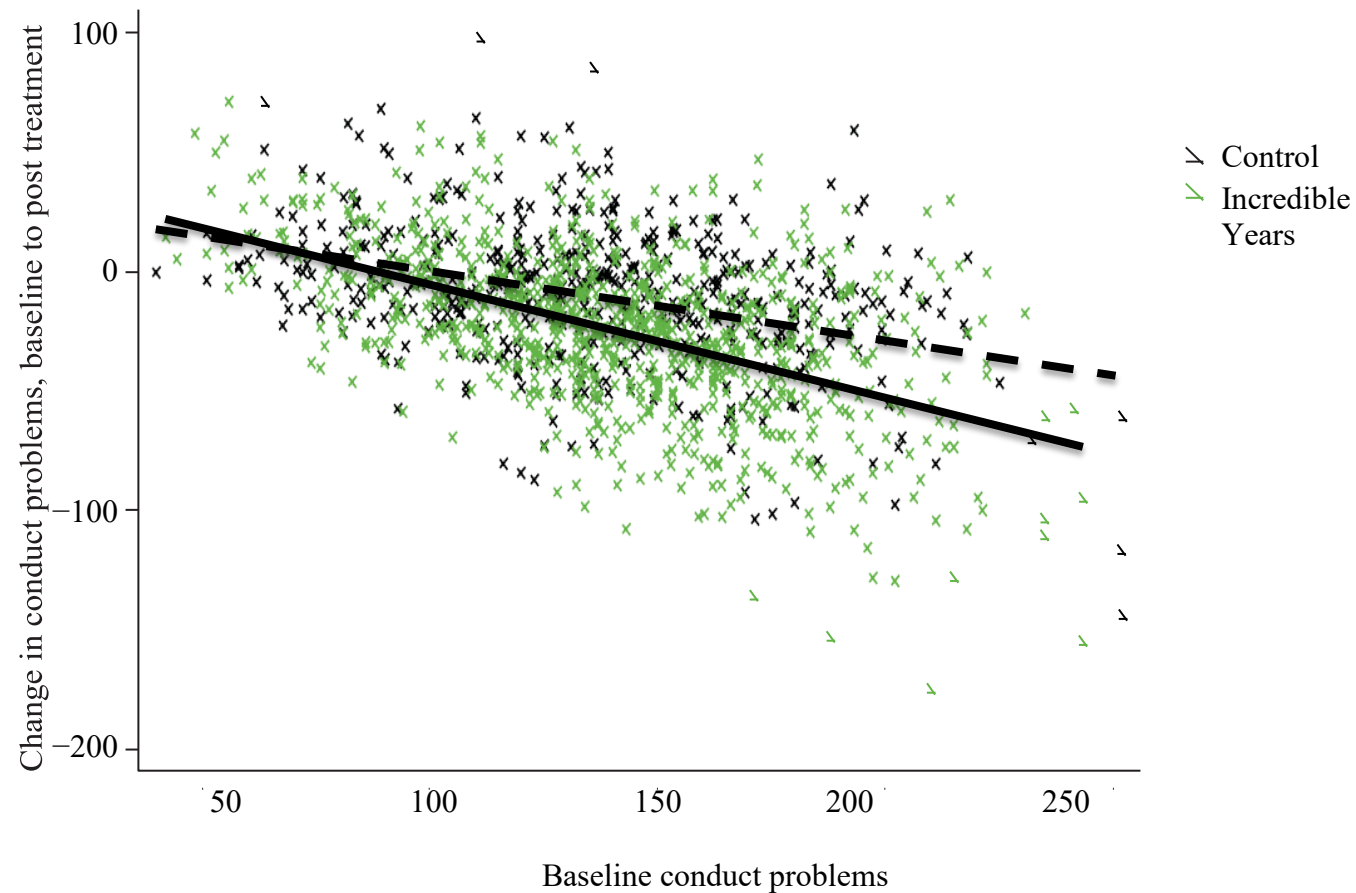


Figure 1. Baseline conduct problems (i.e., Eyberg Child Behavior Inventory scores) moderate the effects of the parenting program on children's conduct problems—children with more severe conduct problems benefit more from the program.

Note. Solid line reflects fitted values Incredible Years; dashed line reflects fitted values control.

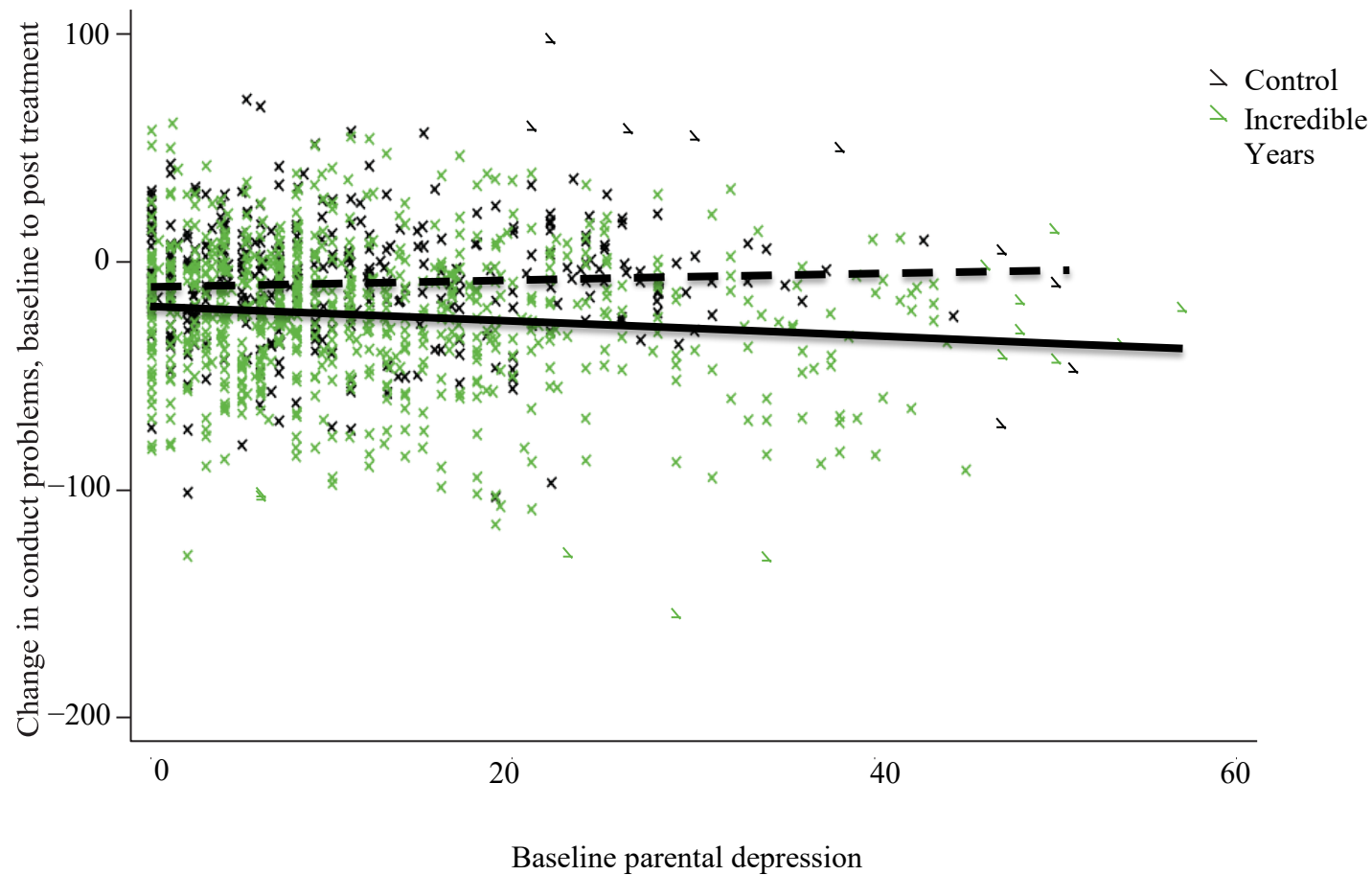


Figure 2. Parental depressive symptoms (i.e., Beck's Depression Inventory scores) moderate the effects of the parenting program on children's conduct problems—children of parents with more depressive symptoms benefit more from the program.

Note. Solid line reflects fitted values Incredible Years; dashed line reflects fitted values control.