Sustained, fade-out or sleeper effects?

A systematic review and meta-analysis of parenting interventions for disruptive child behavior

van Aar, J.; Leijten, P.; Orobio de Castro, Bram; Overbeek, G.

DOI
10.1016/j.cpr.2016.11.006

Publication date
2017

Document Version
Final published version

Published in
Clinical Psychology Review

Clinical Psychology Review 51 (2017) 153–163

Contents lists available at ScienceDirect

Clinical Psychology Review

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

Review

Sustained, fade-out or sleeper effects? A systematic review and meta-analysis of parenting interventions for disruptive child behavior

Jolien van Aar [a,*], Patty Leijten [a,b], Bram Orobio de Castro [c], Geertjan Overbeek [a]

[a] University of Amsterdam, The Netherlands
[b] University of Oxford, United Kingdom
[c] Utrecht University, The Netherlands

HIGHLIGHTS

- We examined follow-up effects of parenting interventions on disruptive child behavior.
- Parenting interventions on average had sustained effects.
- There was much heterogeneity: some trials showed fade-out or sleeper effects.
- None of the moderators tested explained the heterogeneity.

ARTICLE INFO

Article history:
Received 14 May 2016
Received in revised form 4 November 2016
Accepted 11 November 2016
Available online 14 November 2016

Keywords:
Parenting intervention
Disruptive child behavior
Meta-analysis
Follow-up effects
Longer term effects

ABSTRACT

Parenting interventions are known to reduce disruptive child behavior immediately post intervention. But it is largely unknown how reduced disruptive behavior develops in the months and years after the intervention. The present systematic review and multilevel meta-analysis examines whether improvements in disruptive child behavior after parenting intervention are maintained (i.e., sustained effects), fall back (i.e., fade-out effects), or increase further (i.e., sleeper effects). We identified 40 randomized controlled trials with follow-up assessments (up to three years) that generated 91 effect sizes. Mean effect size of post-intervention change was $d = 0.01$, 95% CI [−0.05, 0.07], $p = 0.78$. This lack of change suggests that parenting interventions lead to sustained effects on disruptive behavior. However, there was heterogeneity within and between trials, indicating that some interventions, or interventions under certain circumstances do show fade-out or sleeper effects. None of the moderators tested (i.e., length of follow-up and initial intervention success) explained this heterogeneity. We conclude that parenting interventions generally lead to sustained reductions in disruptive child behavior, at least until three year after intervention. Better understanding is needed of when and why sustainability is stronger in some cases than in others.

Contents

* Corresponding author at: Research Institute of Child Development and Education, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands.
E-mail address: J.vanAar@uva.nl (J. van Aar).

http://dx.doi.org/10.1016/j.cpr.2016.11.006
0272-7358/

1. Introduction

Parenting interventions are one of the most effective strategies to prevent and treat disruptive child behavior, including tantrums, arguing, or rule-breaking (McCart, Priester, Davies, & Azen, 2006; Weisz & Kazdin, 2010). Dozens of systematic reviews and meta-analyses show that parenting interventions reduce disruptive child behavior immediately after intervention (e.g., Menting, Orobio de Castro, & Matthys, 2013; Piquero et al., 2016). Much less is known about the maintenance of intervention effects in the months or years *after* interventions end. A few reviews have examined longer term effects of parenting interventions, by assessing change in disruptive child behavior from pre-intervention until later follow-up (e.g., Lundahl, Risser, & Lovejoy, 2006; Sandler, Schoenfelder, Wolchik, & Mackinnon, 2011). These reviews are critical for estimating the ultimate effectiveness of interventions. However, they do not differentiate between change during the intervention period, and change after the intervention period. Yet, it is specifically the period after interventions that shows whether intervention effects are maintained or change. The present systematic review and meta-analysis addresses this lacuna by reviewing trials on the longer term effects of parenting interventions and by testing the extent to which disruptive child behavior changes between intervention termination and later follow-up.

1.1. Longer term effects of parenting interventions

Behavioral and non-behavioral parenting interventions aim for sustained change in parent-child interactions through improvement in parenting cognitions and/or practices (Kazdin & Weisz, 1998). By supporting parents (e.g., instrumentally or emotionally), changing parents' communication patterns and attitudes towards child rearing, and by teaching them to model positive behavior, to reinforce positive child behavior and to reduce negative behavior, parenting interventions aim to break coercive cycles; behavioral interaction patterns in which disruptive child behavior and negative parenting behavior reinforce each other (Patterson, 1982). The extensively studied immediate benefits of parenting interventions in reducing disruptive child behavior tend to be small to moderate, with smaller effect sizes for prevention trials and larger effect sizes for intervention trials that target clinically severe problems (Lundahl et al., 2006; Nowak & Heinrichs, 2008).

Many parenting intervention evaluation trials use waiting list control designs where families in the control group receive the intervention immediately after posttest assessment (e.g., 53% to 88%, see Lundahl et al., 2006, and Leijten, Raaijmakers, Orobio de Castro, & Matthys, 2013). As a consequence, at later follow-up assessments, the randomized controlled design is no longer intact. Most previous meta-analyses of longer term effects of parenting interventions therefore use effect sizes of within-group differences in the intervention group, as opposed to between-group differences (e.g., Fossum, Handegård, Adolfsen, Vis, & Wynn, 2016; Leijten et al., 2013). Since children in the control group often show a reduction in problem behavior as well (e.g., van Aar, Asscher, Zijsltra, Deković, & Hoffenaar, 2015), within-group effect sizes that do not take any changes in the control group into account tend to be seriously inflated. For example, the review of Lundahl et al. (2006) shows that when longer term effects are based on within-group differences, these are much larger than when based on between-group differences (respectively $d_{w\text{-}g} = -0.87$ versus $d_{b\text{-}g} = -0.21$).

Two previous systematic reviews and meta-analyses have included only trials in which the randomized controlled design was still intact at follow-up. Based on six trials at immediate posttest and three trials at follow-up, Barlow, Smailagic, Ferriter, Bennet, and Jones (2010) found small to moderate immediate and longer term effects of parenting interventions on parent-reported disruptive behavior ($d_{pre\text{-}post} = -0.25$ and $d_{pre\text{-}fu} = -0.38$). Based on 68 trials at immediate posttest and 21 trials at follow-up, Lundahl et al. (2006) also found small to moderate immediate and longer term ($d_{pre\text{-}post} = -0.42$ and $d_{pre\text{-}fu} = -0.21$). The differences between immediate and longer term effect sizes may suggest change in disruptive behavior after the intervention period, but these putative differences between intervention effects immediately post-intervention and at later follow-up were based on small sample sizes and not tested on their significance. Moreover, the suggested post-intervention to follow-up differences were inconsistent: Barlow and colleagues found a stronger effect of parenting interventions at later follow-up than immediately post-intervention; Lundahl and colleagues found a weaker effect of parenting interventions at later follow-up than immediately post-intervention. Therefore, based on this previous work, no conclusions can be drawn on stability versus change in disruptive child behavior after parenting interventions.

1.2. Post-intervention change

Theoretically, three distinct patterns of change in disruptive child behavior after parenting interventions end may seem plausible. First, effects may be sustained once the reciprocal nature of coercive transactions between parent and child has been altered (i.e., sustained effects). Second, effects may fade out, because the modest effects of parenting interventions suggest that these interventions do not fundamentally change interaction patterns in the family (i.e., fade-out effects). Third, effects may even increase, because initially modest effects of changes in family interaction patterns may have a self-reinforcing snowball effect (i.e., sleeper effects). In this systematic review and meta-analysis, we pit against each other these three contrasting

hypotheses about changes in disruptive child behavior after parenting interventions end – that is, from immediate post-intervention to later follow-up (see Fig. 1).

1.2.1. Sustained effects hypothesis

When interventions end, and no further support is provided, the improved parent-child interactions might stay at the same level as where they were at the end of the intervention. Improved parenting behavior may be maintained because the parent is reinforced by experiencing improved child behavior, which is a consequence of their own behavioral change (Rothman, 2000). For example, parents who successfully use limit setting and see how this leads to increased compliance in their child, may experience increased feelings of control and increased self-efficacy as parents. In turn, this might motivate parents to keep using limit setting, and help them to deal effectively with disruptive behavior of their child (Bandura, 1971; Mouton & Roskam, 2015). In the longer term, children who experience more positive and effective parenting may orient themselves less to peers who engage in risky thinking and behavior, which prevents a return to previous, higher levels of disruptive behavior (Goldstein, Davis-Kean, & Eccles, 2005). At the same time, no further decrease in disruptive behavior might be expected when no further support is provided. Therefore, with no further increase or decrease in disruptive behavior, the initial intervention effects are maintained.

1.2.2. Fade-out effects hypothesis

Disruptive child behavior tends to be persistent (Broidy et al., 2003). Although parenting interventions reduce these problems to a certain extent, their effect sizes are small to moderate at best (up to $d = -0.53$ in Furlong et al., 2012) and most children will thus continue to show at least some level of disruptive behavior problems after the intervention. It requires parental strength to keep using newly learned parenting skills in the face of relatively persistent disruptive child behavior. In the absence of continuous support, parents may slowly lose their strength to maintain their improved parenting behavior and fall back to previous (coercive) interaction patterns with their child. This in turn may undo some of the improvements in disruptive child behavior. The use of booster sessions (e.g., Tuning into Kids; Havighurst, Wilson, Harley, Prior, & Kehoe, 2010), the use of telephone sessions to continue minimal contact between parents and therapists after the group sessions (e.g., Triple P group level 4; Sanders, 1999), and the development of intervention programs that are spread out over several years (e.g., Family Check-Up; Dishion et al., 2008) reflect the clinical expectation that families are not always capable of maintaining intervention effects.

1.2.3. Enlarged or "sleeper" effects hypothesis

Parenting intervention is a form of mediational therapy, in which the child is not targeted directly but indirectly, through its parents. It has therefore been suggested that parenting intervention may have sleeper effects on child behavior: a term used to describe the phenomenon of enlarged positive effects at later follow-up, compared to immediately post-intervention (e.g., Barlow et al., 2007; Sofronoff, Jahnel, & Sanders, 2011; Deković et al., 2010). Sleeper effects imply that at least part of the positive intervention effects on child behavior may need more time to materialize (Gray & McCormick, 2005; Kendall, 2006). For example, parents may be unsure about using new parenting skills (e.g., using praise or limit setting) and need some time to implement these well. Later, when parents have practiced and convincingly used their new skills, their behavioral changes may be reinforced by the child's positive response. Similarly, children may need some time to get used to new parenting skills (e.g. receiving "praise" or "time-out") and therefore not improving their behavior directly. In the Incredible Years program, for example, parents are specifically informed to not expect immediate changes, and to even expect some initial resistance from the child (Webster-Stratton, 2006). Therefore, improved parent-child interactions may not have been fully established during the intervention, and may continue to establish after the intervention, leading to gradually increased intervention effects over time.

1.3. Factors that might influence post-intervention change

The fade-out effects and sleeper effects hypotheses indicate gradual change over time in disruptive child behavior after the intervention. Specifically, in the case of fade-out effects, parents may gradually lose their strength to maintain behavioral changes, which would result in stronger fade-out effects over time. Similarly, in the case of sleeper effects, parents and children may continuously reinforce each other's behavior, which would result in stronger sleeper effects over time. By reviewing the evidence base for sustained, sleeper-, or fade-out effects, and by testing whether the time lag (i.e., number of months) between immediate post-intervention and later follow-up is related to these effects, we can identify whether there is indeed a gradual change over time in disruptive child behavior after intervention.

Initial intervention success, defined as the improvement in disruptive child behavior during the intervention period, might be another factor that could affect how child behavior develops after intervention. Families who show more intervention success during the intervention period may have a larger scope for fallback after intervention. Even if families' "larger scope for fallback" does not result in actual fade-out effects, the larger improvements already gained during the intervention period may leave less room for further improvement after the intervention period, resulting in smaller sleeper effects. In contrast, families who show less success during the intervention period may have a smaller

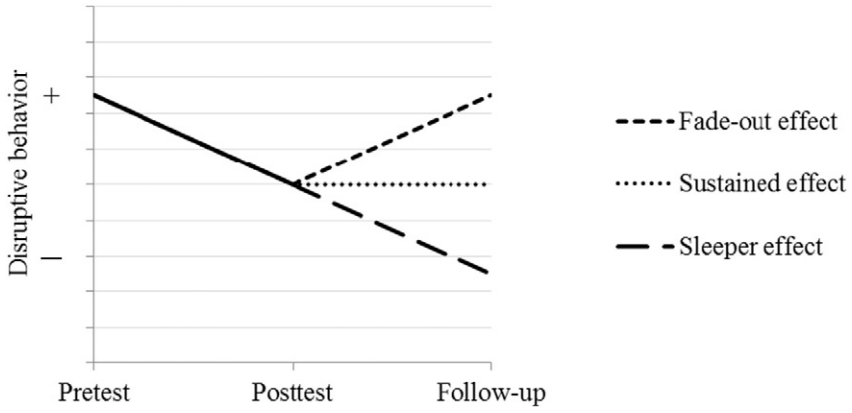

Fig. 1. Graphical display of the sustained effects, fade-out effects and sleeper effects hypotheses.

scope for fallback and a larger scope for improvement left after the intervention period, resulting in smaller fade-out effects or larger sleeper effects. We therefore expect that more initial intervention success is associated with either fade-out or sustained effects, rather than sleeper effects, between intervention termination and later follow-up.

1.4. The present systematic review and meta-analysis

Examining changes in child behavior after the parenting intervention has ended is essential for understanding whether and how parenting intervention effects on child behavior are sustained. We therefore review the body of evidence for longer term effects of parenting interventions and examine whether levels of disruptive child behavior at later follow-up differ from levels of disruptive child behavior immediately after the intervention (i.e., at posttest). Using meta-analysis, we are the first to directly test for patterns of stability versus change in disruptive child behavior after intervention, including only randomized controlled trials with controlled follow-up.

First, we pit three hypotheses against each other: disruptive child behaviors after the intervention show sustained effects, fade-out effects, or sleeper effects. The finding of sustained effects (effect size is zero), follows logically from rejection of both fade-out effects (effect size is positive) and sleeper effects (effect size is negative). Second, we test two putative moderators of changes in child disruptive child behavior after intervention. Because the fade-out and sleeper hypotheses suggest gradual change over time, we expect that longer follow-up periods result in more change in disruptive child behavior than shorter follow-up periods. Because initial intervention success may impact families' scope for fallback or further improvement, we expect that more initial intervention success leads to either larger fade-out effects or smaller sleeper effects than less immediate intervention success.

2. Method

The study protocol of this systematic review and multilevel meta-analysis was registered in PROSPERO (CRD42015020193).

2.1. Information sources

We searched for randomized controlled trials of parenting interventions that targeted child disruptive behavior using the electronic databases PsycINFO, MEDLINE, and ERIC. The initial search was run 6 April 2015, the last update of trials was performed on 4 January 2016. No limit was set on the date of publication. In addition, reference lists of relevant systematic reviews and meta-analyses and experts in the field were consulted. The search terms included all combinations of synonyms of the words parenting (1), child (2), disruptive behavior (3) and follow-up (4). An example of the search terms for the PsycINFO database is provided in Appendix A.

2.2. Trial selection

First, titles of the retrieved trials were reviewed by the first author to determine potential eligibility. Second, abstracts and, if necessary, full texts were reviewed by the first author to determine whether they met the inclusion criteria. Full texts of articles that appeared to meet the criteria were again critically appraised to check if they met inclusion criteria. When there was no full text, authors were requested to deliver those. Duplicate publications of the same data were avoided by juxtaposing author names, intervention comparisons, sample sizes, and pre-intervention outcomes.

Eight criteria were used for inclusion. Trials were included if they 1) reported on a parenting intervention aimed at preventing or reducing child disruptive behavior; 2) performed a pretest, posttest and a follow-up of at least one month; 3) had a randomized intervention and control condition at pretest, posttest and follow-up composed of at least five participants; 4) were aimed at parents or caregivers of children aged 1–15 with a maximum mean age of 12; 5) were not specifically aimed at parents or caregivers of children who were developmentally delayed (e.g. children with intellectual disabilities or problems in the autistic spectrum); 6) involved parenting interventions aimed primarily at parents or caregivers that consisted of at least one face-to-face meeting; 7) used the same parental report of child disruptive behavior across measurement occasions; and 8) were written in English.

2.3. Data extraction

Information was extracted from each included trial on 1) study characteristics (including authors, year and journal of publication, country, type of control condition, drop-out rates and whether or not analyses were based on intention to treat); 2) characteristics of trial participants (including number of parent-child dyads, age, gender, ethnicity and initial behavior problems of the child); 3) type of intervention (including program, dose, delivery (group/individual), booster sessions); and 4) outcome measures of disruptive child behavior (including the instruments used, means and standard deviations of pre-intervention, post-intervention and follow-up scores, length of follow-up occasions; measures of internalizing child behavior or ADHD symptoms not included).

For the meta-analysis, multiple measures of disruptive behavior on multiple follow-up occasions per trial were extracted, generating multiple effect sizes per trial. Where means and/or standard deviations of the pretest, posttest or follow-up scores regarding disruptive behavior were not available in published reports, study authors were contacted to supply missing information. If the reported behavior problem scales included both externalizing and internalizing behavior problems (such as, the total problems scale of the Strengths and Difficulties Questionnaire), we requested scores on disruptive behavior subscale from the authors. Total problems scales were not included in the analysis, because these include both externalizing and internalizing behavior. In cases where two parenting interventions were compared to a control group, we included both parenting interventions as independent comparisons to the control group.

Trials were coded by the first author using a data extraction form for study and sample characteristics, type of intervention and outcome measures. A second coder double coded a random subset of trials (23% of total number of trials). Intraclass correlations ranged between 0.83 and 1.

2.4. Data synthesis for meta-analysis

The primary outcome measure was the standardized difference in change over time between the intervention group and the control group on a continuous variable of child disruptive behavior. For the current research, three effect sizes per trial were calculated using a pretest-posttest controlled design (Morris, 2008): pretest-posttest effect sizes (immediate effects of parenting interventions), pretest-follow-up effect sizes (longer term effects of parenting interventions) and, most important for the aim of the current research, posttest-follow-up effect sizes (follow-up effects of parenting interventions). The estimated effect size was the difference between repeated measures of intervention and control condition, divided by the pooled standard deviation at the first of successive measures (eq. 8, Morris, 2008):

$$d_{ppc} = \left[\frac{(M_{post,T} - M_{pre,T}) - (M_{post,C} - M_{pre,C})}{SD_{pre}}\right]$$

where C_P was a bias adjustment based on sample size (eq. 10, Morris, 2008). Positive effect sizes reflect an increase in disruptive child behavior in the intervention condition compared to the control condition, negative effect sizes reflect a decrease in disruptive child behavior.

The pretest-posttest effect size of parenting interventions on disruptive child behavior was used as the estimate of initial intervention success.

Estimation of the sampling variance of effect sizes requires the mean effect size, standard deviation and test–retest correlation between the pre- and follow-up measure used by the study (eq. 25, Morris, 2008). Test–retest correlations were obtained from each measure's manual or published work about the measure's test-retest reliability (ECBI $r = 0.86$, Robinson, Eyberg & Ross, 1980; CBCL $r = 0.89$, Achenbach & Rescorla, 2001; ITSEA $r = 0.82$, Carter, Briggs-Gowan, Jones & Little, 2003; SDQ $r = 0.64$, Goodman, 2001). If no test-retest correlation was available (which was the case for 30 out of 92 effect sizes), the test–retest correlation was assigned the value of the mean of the other correlations that were available (mean $r = 0.85$).

2.5. Statistical analyses

Because trials reported treatment effects on multiple follow-up occasions and multiple measures, generating multiple effect sizes, the assumption of independency between effect sizes was violated. To account for this dependency in effect sizes, a multilevel approach was used with a Restricted Maximum Likelihood estimation method using the metafor package in R (Viechtbauer, 2010; Assink & Wibbelink, 2016). A three-level random-effects model was estimated which includes the sampling variance for each effect size (Level 1), variance between effect sizes within a study (Level 2) and variance between effect sizes across studies (Level 3). The model generated an overall effect size to indicate whether there is a difference in change in disruptive behavior from immediate posttest to later follow-up between intervention and control condition. An overall effect size from immediate posttest to later follow-up larger than zero would support fade-out effects of parenting interventions, an overall effect size smaller than zero would support sleeper effects of parenting interventions. If the overall effect size would not significantly deviate from zero, this would support sustained effects of parenting interventions.

Next, heterogeneity of effect sizes within (level 2) and across studies (level 3) was estimated by comparing the fit of the original (full) model with the variances randomly estimated to the fit of a model with the variances within or across studies restricted to zero (Assink & Wibbelink, 2016). If the full and restricted model have an equal fit, the variances are not different from zero. We also estimated heterogeneity within and between trials using the 75% rule (Hunter & Schmidt, 1990), which states that there is substantial heterogeneity within and between trials if $<75\%$ of the variance is attributed to sampling variance. The distribution of variance across the three levels was estimated using the Higgins and Thompson method (2002) as proposed by Cheung (2014), to give insight in how much of the variance is a result of sampling variability (level 1), within-study variance (level 2), and between-study variance (level 3).

We subsequently fitted a 3-level mixed-effects model to identify whether the putative moderators length of follow-up and initial intervention success significantly explained within-study variance and between-study variance. Both moderators were included in separate models to test for their significance.

2.6. Risk of bias

To ascertain the validity of eligible randomized trials, the first author determined the adequacy of randomization and concealment of allocation; only trials that had an adequate process of randomization were included. Inequality of pretest scores was accounted for in the effect size. Drop-out rates (i.e., proportion of participants for whom the investigators were not able to ascertain outcomes) were coded per condition per follow-up measurement to provide insight in the reliability of the trial. Intention to treat results (that is, means and standard deviations based on all randomized participants with missing data imputed) were preferred over completers-only results.

Follow-up trials that produced sustained effects or sleeper effects may be more likely to get published than trials without these effects (Dwan et al., 2008). Therefore, we assessed the possibility of publication bias by evaluating a funnel plot and Egger's test of the posttest – follow-up effect sizes for asymmetry. Trim-and-fill procedures were used to examine sustained, fade-out or sleeper effects when publication bias, as standard defined in meta-analysis (Dwan et al., 2008), was removed from the dataset.

3. Results

3.1. Trial selection

Our search yielded 40 randomized controlled trials that evaluated longer term effects of parenting interventions. The searches of PsycInfo, MEDLINE, and ERIC databases provided a total of 6978 citations. Of these, 106 citations remained and were examined in more detail (see Fig. 2 for a flowchart). It appeared that 43 trials did not meet the inclusion criteria as they were not reported in the abstracts: no RCT ($k = 4$), no controlled follow-up ($k = 7$), no pretest performed ($k = 6$), no posttest performed ($k = 6$), age out of range ($k = 4$), no face-to-face meeting ($k = 4$), intervention not primarily targeting parents ($k = 2$), no maternal report of child behavior at all measurement points ($k = 8$), not reported in English ($k = 2$). In addition, some publications reported on an already included trial ($k = 11$), or the full text of the study was not available ($k = 3$). We contacted two authors for full texts (the third author could not be traced), one provided the full text. Twenty-one authors were contacted with a request to provide means and/or standard deviations. Eighteen authors (86%) responded and of these, eleven authors (61% of authors contacted) were able to provide the means and/or standard deviations. Trials for which the required information was not provided ($k = 10$) were excluded.

3.2. Trial characteristics

As can been seen from Table 1, the parenting intervention program that was most frequently evaluated, was Triple P Positive Parenting Program (9 trials), followed by Incredible Years (7 trials) and Parent Management Training Oregon (2 trials). Other parenting interventions were evaluated once. The period of follow-up ranged between one month and three years: 16% had follow-ups between 0 and 6 months, 42% between 6 and 12 months, 35% between 12 and 18 months, and the remaining 7% between 18 and 36 months. In two trials (Havighurst et al., 2010; Mejia, Calam, & Sanders, 2015) parents attended booster sessions before follow-up assessments. Trials varied in the extent to which they focused on the prevention or treatment of disruptive child behavior, ranging from ten trials that included only children with clinical levels of disruptive behavior to three trials that included only children with non-clinical levels of problem behavior.

3.3. Narrative review of post-intervention change

Most trials reported immediate and longer term parenting intervention effects separately. Few trials actually tested whether post-intervention change was significant. Of the five trials that did test for post-intervention change, three trials reported sustained effects on reduced disruptive child behavior. Breitenstein et al. (2012) showed immediate reductions in disruptive child behavior, harsh parenting, and parental feelings of efficacy after participation in the Chicago Parent Program, which stabilized in the 12 months after the intervention. Kirby and Sanders (2014) showed that grandparents and parents reported less disruptive child behavior and more parental confidence, but not improved parenting behavior, immediately after grandparents participated in Grandparent Triple P. These effects stabilized in the six months after the intervention. Finally, Reedtz, Handegård, & Mørch (2011) showed immediate reductions in disruptive child behavior

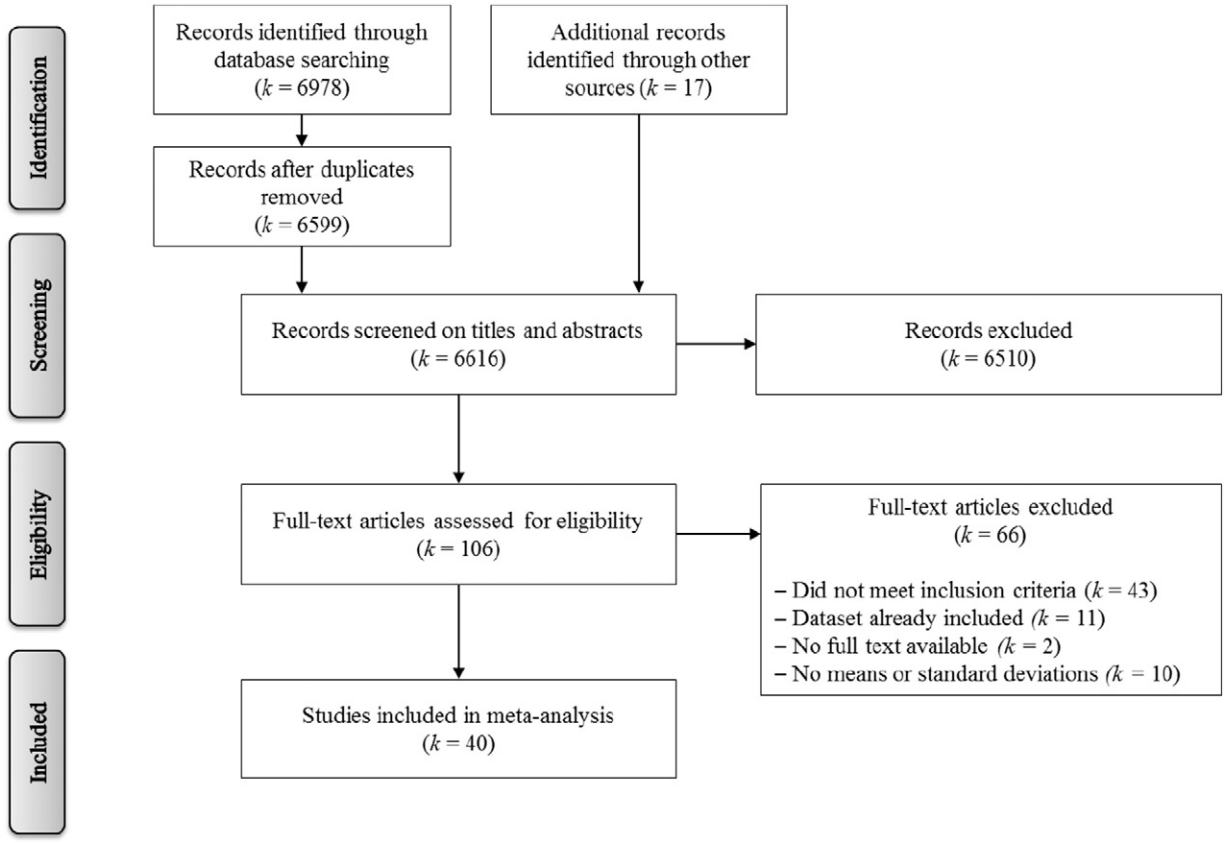

Fig. 2. Flowchart of trial selection.

during intervention period, which stabilized in the year after the intervention, suggesting sustained effects of the shortened version of the Incredible Years parenting program. However, the authors did report their findings as fade-out effects because the immediate effects (i.e., post-intervention group differences) of Incredible Years on disruptive child behavior was not found at later follow-up (i.e., follow-up group differences). Thus, it could be that there was a small fallback in disruptive child behavior in the intervention condition, but not large enough for the significance test of change to identify it. This would be in line with the significant fallback in positive parenting behavior that parents showed. Sustained effects were reported for parental efficacy and harsh parenting behavior. Overall, these three trials demonstrated sustained effects of parenting interventions on child behavior and parental self-efficacy. These trials varied on the examined length of follow-up (6 to 12 months). Because only one trial reported the initial intervention success ($d = -0.82$; Kirby & Sanders, 2014), no inferences could be made about possible influence of initial intervention success for maintenance of intervention success.

Two trials that tested for post-intervention change reported sleeper effects of parenting intervention on reduced disruptive child behavior. Jouriles et al. (2009) showed that disruptive child behavior continued to decrease in the eight months after the Project Support intervention program had ended. These sleeper effects were absent, however, for parenting behavior and parental psychiatric symptoms. Similarly, Somech and Elizur (2012) showed that parents who participated in the Hitkashrut parenting intervention experienced improvement in their children's behavior one year after intervention, when controlled for the improvements immediately after intervention. No follow-up data was reported for measures of parenting behavior. Overall, the examined length of follow-up (8 to 12 months) in these trials was similar to those that reported sustained effects. Initial intervention success was relatively small in one of these trials ($d = -0.17$; Jouriles et al., 2009), but larger in another ($d = -0.70$; Somech & Elizur, 2012). This suggests that post-intervention change may depend in part on differences in immediate intervention success.

Although no significance tests of change were reported, five other trials suggest either fade-out or sleeper effects. Three trials showed a beneficial effect of parenting interventions that disappeared at later follow-up (i.e., fade-out effects; Fabiano et al., 2012; Maguin, Zucker, & Fitzgerald, 1994; Stewart-Brown et al., 2004), and two trials that did not show immediate effect of parenting interventions did show longer term effects (i.e., sleeper effects; Lowell, Carter, Godoy, Paulicin, & Briggs-Gowan, 2011; Niccols, 2009).

The majority of the trials reported similar parenting intervention effects immediately after the intervention and at later follow-up. Importantly, half of these trials ($k = 10$) reported beneficial effects both immediately post-intervention and at later follow-up (e.g., Scott, Sylva, Kallitsoglou, & Ford, 2014), while the other half ($k = 10$) reported neither effects immediately post-intervention, nor at later follow-up (e.g., Forgatch & DeGarmo, 1999). Although no tests of post-intervention change were reported, these trials suggest stability in disruptive child behavior after parenting interventions end.

3.4. Meta-analysis of post-intervention change

Means and standard deviations of child disruptive behavior were available from 40 trials, generating 92 effect sizes. In total, 2955 families were randomized to the intervention condition and 2580 to the control condition. Drop-out rates were generally low and similar for intervention and control conditions (respectively $M = 18\%$ and $M = 19\%$). They ranged from 0% to 36% at immediate posttest ($M = 13\%$, $SD = 7\%$), and from 0% to 43% at later follow-up ($M = 18\%$, $SD = 8\%$). Means and standard deviations based on intention to treat were available for 39% of the effect sizes. Gross, Fogg, and Tucker (1995) reported early drop-outs from the intervention group separately; these data were combined with data from the intervention group to provide an

Table 1
Characteristics of included trials.

Trial	Program	Sample Size IC/CC	Mean age (age range)	Boys (%)	Type of control	Follow-up (months)	Drop-out[a] IC/CC (%)		ITT
							Posttest	Follow-up	
Bodenmann et al. (2008)	Triple P level 4	50/50	6.5 (2.0–12.0)	52	NT	6, 12	–	6/26	No
	CCET	50/50	6.6 (2.0–12.0)	55	NT	6, 12	–	8/26	No
Breitenstein et al. (2012)	Chicago parent program	267/283	2.8 (2.0–4.0)	54	NT	6, 12	–	13/11	No
Brock et al. (2015)	Child's game	93/91	2.5 (2.0–3.5)	52	CAU	6	5/12	10/14	No
Cowan et al. (2009)	Couples program	129/124	2.3 (0.0–7.0)	–	MC	9	28/19	26/21	No
Fabiano et al. (2012)	Coaches	28/27	8.5 (6.0–12.0)	87	NT	1	7/15	18/15	No
Forgatch and DeGarmo (1999)	Parenting through change	153/85	7.8 (6.1–10.4)	100	NT	6	–	18/18	No
Frank et al. (2015)	Triple P, level 4	23/19	5.6 (3.0–8.0)	69	NT	6	9/0	9/5	Yes
Gross et al. (1995)	Parent training program	11/6	2.0 (2.0–3.0)	83	NT	3	9/0	9/0	No
Gross et al. (2003)	Incredible Years	75/59	– (2.0–3.0)	–	NT	6, 12	16	21	No
Hahlweg et al. 2010	Triple P, level 4	158/61	4.5 (2.6–6.0)	51.0	NT	12, 24	23/0	23/2	Yes
Hanish et al. (2010)	PEP	60/65	4.2 (3.0–7.0)	–	NT	6, 18, 30	–	42/43	No
Havighurst et al. (2010)	Tuning in to kids	106/110	4.7 (3.8–5.7)	52.3	NT	6	20/13	17/7	No
Jouriles et al. (2001)	Multicomponent family intervention	18/18	5.7 (4.0–9.0)	72	MC	4, 8	–	–	–
Jouriles et al. (2009)	Project support	32/34	– (4.0–9.0)	50	MC	4, 8, 12	–	16/15	Yes
Kim et al. (2008)	Incredible Years	12/21	6.0 (3.0–8.0)	63	NT	12	5/25	19/42	No
Kirby and Sanders (2014)	Grandparent Triple P	28/26	4.9 (2.0–9.0)	61	CAU	6	8/4	14/4	Yes
Kjøbli and Bjørnebekk (2013)	Brief parent training	108/108	7.3 (3.0–12.0)	68	CAU	6	12/15	17/23	Yes
Kjøbli, Hukkelberg, and Ogden (2013)	PMTO	72/65	8.6 (2.0–12.0)	64	NT	6	8/11	8/9	Yes
Kleefman et al. (2011)	Triple P, stepping stones	111/98	9.8 (5.0–12.0)	58	CAU	6	8/4	10/7	Yes
Lavigne et al. (2008)	Incredible Years	86/31	4.6 (3.0–6.1)	53	MC	12	26/13	23/0	No
Lowell et al. (2011)	Child first	78/79	1.5 (0.5–3.0)	44	CAU	6	18/15	26/25	Yes
Magen and Rose, (1994)	Problem-solving training	18/19	7.0 (5.0–11.0)	70	NT	3	–	–	No
	Behavioral skills training	19/19	7.0 (5.0–11.0)	70	NT	3	–	–	No
Maguin et al. (1994)	Program based on OSLC protocol	29/23	4.4 (3.0–6.0)	100	NT	6	3/0	0/4	No
Malti et al. (2011)	Triple P, level 4	339/356	7.5 (–)	52	NT	12, 36	4/4	5/4	Yes
McCabe et al. (2012)	GANA	21/18	4.4 (3.0–7.0)	71	CAU	16	5/11	5/28	Yes
	PCIT	19/18	4.4 (3.0–7.0)	71	CAU	16	5/11	21/28	Yes
Mejia et al. (2015)	Triple P, level 3	54/54	8.5 (3.0–12.0)	70	NT	6	11/15	35/24	Yes
Niccols (2009)	CWTB	49/30	2.0 (1.0–3.0)	60	NT	1	8/3	8/13	No
Perrin et al. (2014)	Incredible Years, adapted	89/61	2.8 (1.8–3.5)	62	NT	12	–	19/18	No
Reedtz et al. (2011)	Incredible Years, adapted	89/97	3.9 (2.0–8.0)	59	NT	12	25/46	25/53	No
Rushton et al. (2010)	Cognitive behavioral approach	19/18	6.6 (1.5–10.0)	46	CAU	6	0/0	0/0	Yes
Scott et al. (2014)	Incredible Years	55/54	6.1 (5.0–7.0)	56	CAU	12	19/13	26/20	–
Sheeber and Johnson, (1994)	Temperament-based parent-training	20/20	4.0 (3.0–5.0)	60	NT	2	5	12	No
Sigmarsdóttir et al. (2013)	PMTO	51/51	8.0 (5.0–12.0)	73	CAU	9	2/8	10/14	No
Somech and Elizur, (2012)	Hitkashrut	140/69	4.1 (2.7–5.3)	78	MC	12	11/17	31/58	Yes
Sonuga-Barke et al. (2001)	Parent training	30/20	– (3.0–4.0)	62	NT	3	–	9	Yes
	Parent counseling and support	28/20	– (3.0–4.0)	62	NT	3	–	9	Yes
Spijkers et al. (2013)	Triple P, level 3	42/39	10.6 (9.0–11.0)	56	CAU	6, 12	24/15	29/35	No
Stewart-Brown et al. (2004)	Incredible Years	60/56	4.6 (2.0–8.0)	–	NT	6, 12	23/11	28/23	No
Thompson et al. (2009)	New forest parenting program	21/20	4.3 (2.5–6.4)	76	CAU	2	19/22	10/28	No
Westrupp et al., (2015)	Triple P, level 4	38/38	9.0 (4.0–12.0)	51	CAU	12	24/18	16/34	No
Williamson et al. (2014)	Mothers to mothers	113/81	9.47 (7.0–12.0)	45	NT	6	6/6	12/23	No

Note. Dashes indicate that no data was reported, IC = intervention condition, CC = control condition, ITT = intention to treat, CCET = couples coping enhancement training, PEP = prevention program for externalizing problem behavior, PMTO = parent management training, the Oregon model, OSCL = Oregon social learning center, GANA = Guiando a Ninos Activos, PCIT = parent child interaction therapy, CWTB = COPEing with Toddler behavior, NT = no treatment, MC = minimal contact, CAU = care as usual.

[a] Drop-out rates are reported for the last follow-up occasion. If drop-out rates were only available for the total sample, one rate is reported.

intention to treat estimate for this trial. Effect sizes for the other trials were based on the means and standard deviations of the subset of participants who were retained until later follow-up.

3.4.1. Results of individual trials

The individual effect sizes per trial are provided in Appendix B, a graphical display of the effect sizes over time is shown in Fig. 3. Effect sizes scattered around zero, without a specific trend over time. There was one extreme value ($d = 1.24$), which was based on father-reported follow-up means and standard deviations. Close consideration revealed that the relatively high means and standard deviations at follow-up were exactly the same as the mother-reported means and standard deviations at follow-up (for whom the scores were relatively stable). Because of outdated data files, we could not confirm that these data were correct. Therefore, we decided to omit this effect size from further analyses, resulting in 40 trials and 91 effect sizes that ranged from $d = -0.65$ to $d = 0.65$, with 16% of the effect sizes indicating significant fade-out effects and 12% of the effect sizes indicating significant sleeper effects (95% CI does not include zero).

3.4.2. Synthesis of results

The overall effect size of parenting interventions, relative to the control conditions, on reduced disruptive child behavior between pretest and immediate posttest was significant and small to moderate ($d = -0.32$, 95% CI [−0.40, −0.23], $p < 0.001$). The overall effect size on reduced disruptive behavior from pretest to follow-up was also small to moderate ($d = -0.31$, 95% CI [−0.41, −0.21], $p < 0.001$). More importantly for the purpose of this review, the overall effect size from posttest to follow-up was non-significant ($d = 0.01$, 95% CI [−0.05, 0.07], $p = 0.78$). This demonstrates stability in the effects of

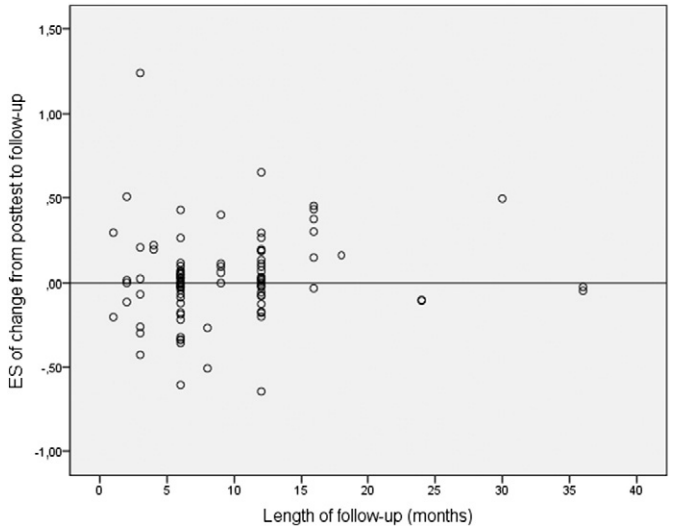

Fig. 3. Scatter plot with on the X-axis the number of months between immediate posttest and later follow-up and on the Y-axis the effect sizes of change in disruptive child behavior from immediate posttest to later follow-up. The bullet with the largest effect size (d = 1.24) was considered an outlier.

parenting interventions on disruptive child behavior between immediate post-intervention and later follow-up. Parenting interventions lead to sustained effects on disruptive child behavior.

However, significant heterogeneity was estimated between effect sizes from post-intervention to follow-up within trials ($\chi^2(1) = 49.36$, $p < 0.001$), and across trials ($\chi^2(1) = 9.90$, $p = 0.002$). This means that the variance between effect sizes within and across trials is larger than zero. The distribution of variances across levels gives insight in how much of the variance could be attributed to sampling variance, within-study variance and between-study variance. The percentage of sampling variability was 18% (level 1), the percentage of variance in effect sizes that could be attributed to differences within trials was 38% (level 2), and the percentage of variance in effect sizes that could be attributed to differences between trials was 44% (level 3). The relatively low percentage of sampling variability (<75%) supports that there is more variance between effect sizes than can be expected from chance. This suggests that there could be within-study (e.g., length of follow-up occasion) and between-study characteristics (e.g., initial intervention success) that influence the sustainability of intervention effects.

3.4.3. Moderator analyses

We tested two factors that might influence the extent to which disruptive child behavior changes after the intervention. First, in order to identify gradual change over time, we tested whether the length of follow-up was related to the estimate of change in disruptive child behavior after parenting interventions. The Omnibus-test revealed that length of follow-up was not significantly related to the estimate of change in disruptive behavior from immediate posttest to later follow-up ($F(1, 89) = 0.06$, $p = 0.815$). The absence of gradual change over time supports our finding that parenting interventions generally lead to sustained change on disruptive child behavior. Second, initial intervention success was not related to the overall effect ($F(1, 89) = 1.98$, $p = 0.163$), indicating that the improvements made during the intervention period were not related to change in disruptive child behavior after the intervention.

Since there was substantial heterogeneity that was not explained by differences in length of follow-up and initial intervention success, we tested whether any of the available participant characteristics (child's age and gender, ethnicity, initial severity of behavior problems), intervention characteristics (intervention program, number of sessions, delivery format, use of booster sessions) and design characteristics (control condition, type of informant, type of instrument) were related to post-intervention change in disruptive child behavior. None of them were. All models showed a sustained effect of parenting interventions on disruptive child behavior (see Appendix C for exact results).

3.4.4. Robustness of findings

Fig. 4 shows the funnel plot with the effect sizes of parenting interventions on disruptive child behavior after the intervention on the X-axis and the standard error of the effects on the Y-axis. An effect size of zero means that there is no change in children's disruptive behavior from posttest to follow-up, it indicates sustained effects of parenting interventions. Effect sizes at the right side of zero indicate fade-out effects, effect size on the left side of zero indicate sleeper effects. Asymmetry in the plot with an absence of fade-out effects might suggest publication bias. The plot appears symmetrical, Egger's regression test showed no significant asymmetry ($t(89) = 7.70$, $p = 0.484$). Also, the trim-and-fill procedure revealed no additional effect sizes. This means that there are no reasons to expect that the results are influenced by publication bias.

Sensitivity analyses showed that the effect was the same when other assumptions were made or other criteria were used to include trials. Inclusion of the outlier did not result in different follow-up effects. Exclusion of trials with high (>30%) drop-out rates, or exclusion of effect sizes that were not based on intention to treat did not lead to different follow-up effects. Inclusion of solely traditional or standard parent training programs did not lead to different follow-up effects. Also, extending the minimum period of follow-up to six months (the minimum length of follow-up according to standards of evidence as formulated by the Society for Prevention Research, Flay et al., 2005) did not lead to different follow-up effects. All subsets of trials showed sustained effects of parenting interventions on disruptive child behavior (see Appendix C for exact results).

4. Discussion

We systematically reviewed the literature on sustained, fade-out or sleeper effects of parenting interventions for disruptive child behavior. There was much variation in the extent to which trials examined post-intervention changes in disruptive child behavior and few trials actually tested whether these changes were significant. In a multilevel meta-

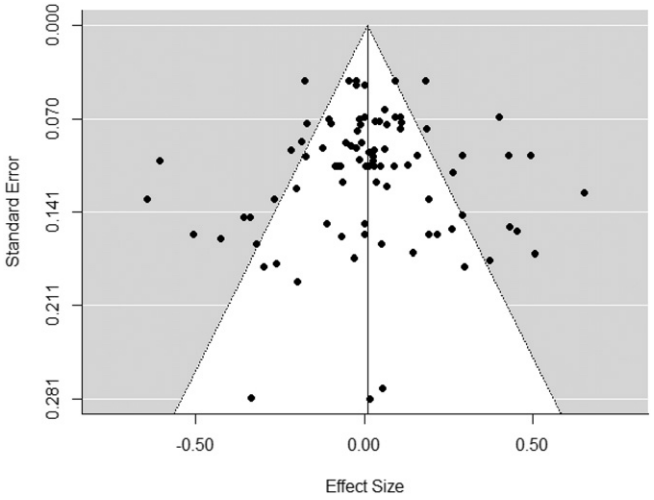

Fig. 4. Funnel plot with on the X-axis the posttest – follow-up effect sizes and on the Y-axis the standard errors (a higher standard error reflects a lower sample size).

analysis, we tested for sustained, fade-out or sleeper effects of parenting interventions for disruptive child behavior. Our results of 40 randomized controlled trials (91 effect sizes) demonstrates sustained effects of parenting interventions on disruptive child behavior; the overall effect size of change *after* intervention (posttest – later follow-up) was close to zero and non-significant. There was no gradual change in disruptive child behavior after intervention and initial intervention success did not influence post-intervention change. This rejects the fade-out and sleeper effects hypotheses: children's disruptive behavior after intervention, at least until three years follow-up, remains stable.

Our finding that the effects of parenting interventions are sustained for months or years after the end of intervention is in line with the small to moderate effects that previous meta-analyses have found for both the short and longer term effects of parenting interventions (Barlow et al., 2010; Lundahl et al., 2006). Our finding suggests that parents keep using their new parenting skills, also after parenting interventions end, at least to the extent that parents are able to prevent children from returning to more disruptive behavior. Parents' use of these new skills might be reinforced by improvements in their children's behavior (Rothman, 2000). Indeed, mediation studies on the effects of parenting interventions on disruptive behavior have identified that reduced disruptive child behavior as a result of parenting intervention has beneficial effects for parents (DeGarmo, Patterson, & Forgatch, 2004). This suggests that reciprocal effects between parent and child behavior following parenting intervention might lead to sustained improvement in parent-child interactions.

There was substantial variance between trials in the extent to which disruptive child behavior changed after parenting interventions. Effect sizes of change between intervention termination and later follow-up in disruptive child behavior ranged from $d = -0.65$ to $d = 0.65$. More specifically, 16% of the individual effect sizes indicated significant fade-out effects and 12% indicated significant sleeper effects of parenting interventions. This means that the overall sustained effect of parenting interventions should be interpreted with caution; under certain circumstances parenting interventions may lead to fade-out effects while under other circumstances parenting interventions may lead to the opposite –sleeper– effects. None of the included moderators were able to account for the differences in findings from individual trials. Neither length of follow-up, nor initial interventions success, influenced change in disruptive child behavior after intervention. Moreover, none of the additional moderators that could be extracted on the level of participant (e.g., initial severity of problem behavior), intervention (e.g., specific intervention program), and design characteristics (e.g., type of control condition) influenced change in disruptive child behavior after intervention.

Other potentially relevant moderators, for which the information could not be extracted, may influence change in disruptive child behavior after intervention. For example, it might be more difficult for parents to keep using their newly learned skills in the face of contextual difficulties such as low SES and neighborhood dangerousness (Trentacosta et al., 2008). The stressors that may accompany these difficulties (e.g., maternal depressive symptoms, Gardner, Hutchings, Bywater, & Whitaker, 2010) may make it harder for families to maintain, or further increase, positive change, resulting in fade-out effects rather than sleeper effects. Second and related, feelings of parental efficacy might influence change in disruptive child behavior after intervention. Increased feelings of parental efficacy may help parents to deal more effectively with present and future disruptive behavior patterns that require new parenting strategies, resulting in sleeper rather than fade-out effects (Bandura, 1971). Finally, parental satisfaction with initial intervention success might influence change in disruptive child behavior after intervention. If improvements in child behavior meet parents' expectations they might be more motivated to maintain their improved parenting skills, resulting in sustained or sleeper rather than fade-out effects (Rothman, 2000). Future research should therefore examine these and other putative predictors of change in disruptive child behavior after parenting interventions.

For clinical practice, our conclusion is relatively straight-forward: neither further improvements, nor worsening of disruptive child behavior, should be expected in the months or years after parenting interventions end. This suggests that if families do not respond to the intervention during the intervention period, a different approach is needed, rather than awaiting possible sleeper effects. At the same time, because generally no fade-out effects should be expected, booster sessions to prevent fallback may not be as important as sometimes suggested (e.g., Eyberg, Edwards, Boggs, & Foote, 1998). Yet, our findings also showed much variation in post-intervention change across trials and in some cases sleeper and fade-out effects did occur. Identification of possible subgroups of families that show different patterns of change is essential, because these families may need different types of support. For example, families who show sleeper effects might need more time to change, rather than a different approach, whereas families who show fade-out effects might benefit from additional booster sessions to prevent fallback. Knowledge on family characteristics that might predict patterns of change can help tailoring parenting interventions to families' individual needs.

Several limitations of our review should be taken into consideration. First, longitudinal trials often suffer from drop-out. In the included trials drop-out rates ranged between 0% and 43% ($M = 18\%$). As only 39% of the effect sizes could be calculated with means and standard deviations based on an intention to treat approach, a relatively large proportion of the analyzed data was incomplete. Although we found similar effects for trials that did and did not provide means and standard deviations based on intention to treat, we cannot rule out that selective drop-out occurred. Selective drop-out can induce bias, for example when families in the experimental condition are more likely to drop out when they have more severe problems. Second, we relied mainly on parent-reported disruptive child behavior. Parent reports can also induce bias, as parents are aware that they participate in a parenting intervention (Sonuga-Barke et al., 2013). The time and effort they devote to the parenting intervention may cause them to overemphasize its beneficial effects. Only in very few cases there were additional teacher reports available and included. Finally, meta-analyses are strong in estimating effects on study level, but cannot draw any conclusions on individual level. Pooled analyses are needed improve the knowledge of individual-level characteristics that may influence post-intervention change.

Despite these limitations, this study was the first to systematically review and meta-analyze change versus stability in parenting intervention effects on child disruptive behavior after interventions. While other meta-analyses focused on change between pre-intervention and later follow-up (long term effects, e.g., Lundahl et al., 2006), we focused on change between intervention termination and later follow-up. This allowed us to test whether initial reductions in disruptive child behavior after parenting interventions are maintained or change in the months or years after intervention. We included only randomized controlled trials that were still intact at later follow-up, in contrast to many other reviews and meta-analyses that also included follow-up data for which controlled data was not available due to waitlist assignment (e.g., Fossum et al., 2016; Leijten et al., 2013). By limiting our sample of trials to those with randomized assignment that allowed comparison of changes in the intervention group to changes in the control group, we have sought to estimate the change in disruptive child behavior as accurately as possible.

5. Conclusion

The substantial variance across trials in changes in disruptive child behavior in the months or years after parenting interventions end does not allow us to conclude straight off that parenting

interventions lead to sustained effects. We therefore mainly conclude that there is a need to improve our understanding of the circumstances under which sustained, fade-out and sleeper effects occur. This knowledge is vital for tailoring parenting interventions to the needs of individual families, in order to strengthen intervention effects. That said, our results show that disruptive child behavior generally does not change much after interventions end. We therefore tentatively conclude that "what you see is what you get:" effects of parenting interventions on disruptive child behavior tend to sustain, rather than change, up to at least three years after intervention.

Supplementary data

Supplementary data (Appendix A, Appendix B and Appendix C) to this article can be found online at http://dx.doi.org/10.1016/j.cpr.2016.11.006.

References

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA school-age forms & profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. http://dx.doi.org/10.20982/tqmp.12.3.p154.

Bandura, A. (1971). *Social learning theory.* New York City, NY: General Learning Press.

Barlow, J., Davis, H., McIntosh, E., Jarrett, P., Mockford, C., & Stewart-Brown, S. (2007). Role of home visiting in improving parenting and health in families at risk of abuse and neglect: results of a multicentre ranomised controlled trial and economic evaluation. *Archives of Disease in Childhood*, *92*, 229–233. http://dx.doi.org/10.1136/adc.2006.095117.

Barlow, J., Smailagic, N., Ferriter, M., Bennet, C., & Jones, H. (2010). Group-based parent-training programmes for improving emotional and behavioral adjustment in children from birth to three years old. *Cochrane Database of Systematic Reviews*, *3*, CD003680. http://dx.doi.org/10.1002/14651858.CD003680.pub2.

Bodenmann, G., Cina, A., Ledermann, T., & Sanders, M. R. (2008). The efficacy of the triple P-positive parenting program in improving parenting and child behavior: A comparison with two other treatment conditions. Behaviour Research &amp. *Therapy*, *46*(4), 411–427. http://dx.doi.org/10.1016/j.brat.2008.01.001.

Breitenstein, S. M., Gross, D., Fogg, L., Ridge, A., Garvey, C., Julion, W., & Tucker, S. (2012). The Chicago parent program: Comparing 1-year outcomes for African American and Latino parents of young children. *Research in Nursing & Health*, *35*(5), 475–489. http://dx.doi.org/10.1002/nur.21489.

Brock, R. L., Kochanska, G., O'Hara, M. W., & Grekin, R. S. (2015). Life satisfaction moderates the effectiveness of a play-based parenting intervention in low-income mothers and toddlers. *Journal of Abnormal Child Psychology*, *43*, 1283–1294. http://dx.doi.org/10.1007/s10802-015-0014-y.

Broidy, L. M., Tremblay, R. E., Brame, B., Fergusson, D. M., Horwood, J. L., Laird, R., et al. , & Vitaro, F.(2003). Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: A six-site, cross-national study. *Developmental Psycholoy*, *39*(2), 222–245. http://dx.doi.org/10.1037/0012-1649.39.2.222.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The infant-toddler social and emotional assessment (ITSEA): Factor structure, reliability, and validity. *Journal of abnormal Child Psychology*, *31*(5), 495–514. http://dx.doi.org/10.1023/A:1025449031360.

Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*, 211–229. http://dx.doi.org/10.1037/a0032968.

Cowan, P. A., Cowan, C. P., Pruett, M. K., Pruett, K., & Wong, J. J. (2009). Promoting fathers' engagement with children: Preventive interventions for low-income families. *Journal of Marriage and Family*, *71*(3), 663–679. http://dx.doi.org/10.1111/j.1741-3737.2009.00625.x.

DeGarmo, D. S., Patterson, G. R., & Forgatch, M. S. (2004). How do out-comes in a specified parent training intervention maintain or wane over time? *Prevention Science*, *5*(2), 73–89. http://dx.doi.org/10.1023/B:PREV.0000023078.30191.e0.

Deković, M., Asscher, J. J., Hermanns, J., Reitz, E., Prinzie, P., & van den Akker, A. L. (2010). Tracing changes in families who participated in the Home-Start parenting program: Parental sense of competence as mechanism of change. *Preventive Sciences*, *11*, 263–274. http://dx.doi.org/10.1007/s11121-009-0166-5.

Dishion, T. J., Connell, A., Weaver, C., Shaw, D., Gardner, F., & Wilson, M. (2008). The family check-up with high-risk indigent families: preventing problem behavior by increasing parents' positive behavior support in early childhood. *Child Development*, *79*(5), 1395–1414. http://dx.doi.org/10.1111/j.1467-8624.2008.01195.x.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A., Cronin, E., et al. , & Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, *3*(8), e3081. http://dx.doi.org/10.1371/journal.pone.0003081.

Eyberg, S. M., Edwards, D., Boggs, S. R., & Foote, R. (1998). Maintaining the treatment effects of parent training: The role of booster sessions and other maintenance strategies. *Clinical Psychology: Science and Practice*, *5*(4), 544–554. http://dx.doi.org/10.1111/j.1468-2850.1998.tb00173.x.

Fabiano, G. A., Pelham, W. E., Cunningham, C., Yu, J., Gangloff, B., Buck, M., ... Gera, S. (2012). A waitlist-controlled trial of behavioral parent training for fathers of children with ADHD. *Journal of Clinical Child and Adolescent Psychology*, *41*(3), 337–345. http://dx.doi.org/10.1080/15374416.2012.654464.

Flay, B. R., Biglan, A., Boruch, R. F., González Castro, F., Gottfredson, D., Kellam, S., et al. , & Ji, P.(2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, *6*(3), 151–175. http://dx.doi.org/10.1007/s11121-005-5553-y.

Forgatch, M. S., & DeGarmo, D. S. (1999). Parenting through change: An effective prevention program for single mothers. *Journal of Consulting & Clinical Psychology*, *67*(5), 711–724. http://dx.doi.org/10.1037/0022-006X.67.5.711.

Fossum, S., Handegård, B. H., Adolfsen, F., Vis, S. A., & Wynn, R. (2016). A meta-analysis of long-term outpatient treatment effects for children and adolescents with conduct problems. *Journal of Child and Family Studies*, *25*(1), 15–29. http://dx.doi.org/10.1007/s10826-015-0221-8.

Frank, T. J., Keown, L. J., & Sanders, M. R. (2015). Enhancing father engagement and interparental teamwork in an evidence-based parenting intervention: A randomized-controlled trial of outcomes and processes. *Behavior Therapy*, *46*, 749–763. http://dx.doi.org/10.1016/j.beth.2015.05.008.

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews*, *2*, CD008225. http://dx.doi.org/10.1002/14651858.CD008225.pub2.

Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child and Adolescent Psychology*, *39*(4), 568–580. http://dx.doi.org/10.1080/15374416.2010.486315.

Goldstein, S. E., Davis-Kean, P. E., & Eccles, J. S. (2005). Parents, peers, and problem behavior: A longitudinal investigation of the impact of relationship perceptions and characteristics on the development of adolescent problem behavior. *Developmental Psychology*, *41*(2), 401–413. http://dx.doi.org/10.1037/0012-1649.41.2.401.

Goodman, R. (2001). Psychometric properties of the strenghts and difficulties questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(11), 1337–1345. http://dx.doi.org/10.1097/00004583-200111000-00015.

Gray, R., & McCormick, M. C. (2005). Early childhood intervention programs in the US: Recent advances and future recommendations. *The Journal of Primary Prevention*, *26*(3), 259–275. http://dx.doi.org/10.1007/s10935-005-3600-x.

Gross, D., Fogg, L., & Tucker, S. (1995). The efficacy of parent training for promoting positive parent–toddler relationships. *Research in Nursing & Health*, *18*(6), 489–499. http://dx.doi.org/10.1002/nur.4770180605.

Gross, D., Fogg, L., Webster-Stratton, C., Garvey, C., Julion, W., & Grady, J. (2003). Parent training of toddlers in day care in low-income urban communities. *Journal of Consulting & Clinical Psychology*, *71*(2), 261–278. http://dx.doi.org/10.1037/0022-006X.71.2.261.

Hahlweg, K., Heinrichs, N., Kuschel, A., Bertram, H., & Naumann, S. (2010). Long-term outcome of a randomized controlled universal prevention trial through a positive parenting program: is it worth the effort? *Child and Adolescent Psychiatry and Mental Health*, *4*, 14. http://dx.doi.org/10.1186/1753-2000-4-14.

Hanisch, C., Hautmann, C., Eichelberger, I., Pluck, J., & Dopfner, M. (2010). Long-term clinical significance of the prevention programme for externalizing problem behaviour (PEP). *Verhaltenstherapie*, *20*(4), 265–273. http://dx.doi.org/10.1159/000322044.

Havighurst, S. S., Wilson, K. R., Harley, A. E., Prior, M. R., & Kehoe, C. (2010). Tuning in to kids: Improving emotion socialization practices in parents of preschool children -Findings from a community trial. *Journal of Child Psychology and Psychiatry*, *51*(12), 1342–1350. http://dx.doi.org/10.1111/j.1469-7610.2010.02303.x.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. http://dx.doi.org/10.1002/sim.1186.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Jouriles, E. N., McDonald, R., Spiller, L., Norwood, W. D., Swank, P. R., Stephens, N., ... Buzy, W. M. (2001). Reducing conduct problems among children of battered women. *Journal of Consulting and Clinical Psychology*, *69*(5), 774–785. http://dx.doi.org/10.1037//0022.0006X.69.5.774.

Jouriles, E. N., McDonald, R., Rosenfield, D., Stephens, N., Corbitt-Shindler, D., & Miller, P. C. (2009). Reducing conduct problems among children exposed to intimate partner violence: A randomized clinical trial examining effects of project support. *Journal of Consulting and Clinical Psychology*, *77*(4), 705–717. http://dx.doi.org/10.1037/a0015994.

Kazdin, A. E., & Weisz, J. R. (1998). Identifying and developing empirically supported child and adolescent treatments. *Journal of Consulting and Clinical Psychology*, *66*(1), 19–36. http://dx.doi.org/10.1037/0022-006X.66.1.19.

Kendall, P. C. (2006). *Child and adolescent therapy: Cognitive-behavioral procedures* (3rd ed.). New York: The Guilford Press.

Kim, E., Cain, K. C., & Webster-Stratton, C. (2008). The preliminary effect of a parenting program for Korean American mothers: A randomized controlled experimental study. *International Journal of Nursing Studies*, *45*(9), 1261–1273. http://dx.doi.org/10.1016/j.ijnurstu.2007.10.002.

Kirby, J. N., & Sanders, M. R. (2014). A randomized controlled trial evaluating a parenting program designed specifically for grandparents. *Behaviour Research and Therapy*, *52*, 35–44. http://dx.doi.org/10.1016/j.brat.2013.11.002.

Kjøbli, J., & Bjørnebekk, G. (2013). A randomized effectiveness trial of brief parent training: Six-month follow-up. *Research on Social Work Practice*, *23*(6), 603–610 612 10.1177/1049731513492860

Kjøbli, J., Hukkelberg, S., & Ogden, T. (2013). A randomized trial of group parent training: Reducing child conduct problems in real-world settings. Behaviour Research &amp. *Therapy*, *51*(3), 113–121. http://dx.doi.org/10.1016/j.brat.2012.11.006.

Kleefman, M., Jansen, D. E., & Reijneveld, S. A. (2011). The effectiveness of Stepping Stones Triple P: the design of a randomised controlled trial on a parenting programme regarding children with mild intellectual disability and psychosocial problems versus care as usual. *BioMed Central public health*, *11*, 676. http://dx.doi.org/10.1186/1471-2458-11-676.

Lavigne, J. V., Lebailly, S. A., Gouze, K. R., Cicchetti, C., Pochyly, J., Arend, R., ... Binns, H. J. (2008). Treating oppositional defiant disorder in primary care: a comparison of three models. *Journal of Pediatric Psychology*, *33*(5), 449–461. http://dx.doi.org/10.1093/jpepsy/jsm074.

Leijten, P., Raaijmakers, M. A. J., Orobio de Castro, B., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. *Journal of Clinical Child & Adolescent Psychology*, *42*(3), 384–392. http://dx.doi.org/10.1080/15374416.2013.769169.

Lowell, D. I., Carter, A. S., Godoy, L., Paulicin, B., & Briggs-Gowan, M. J. (2011). A randomized controlled trial of child FIRST: A comprehensive home-based intervention translating research into early childhood practice. *Child Development*, *82*(1), 193–208. http://dx.doi.org/10.1111/j.1467-8624.2010.01550.x.

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

Magen, R. H., & Rose, S. D. (1994). Parents in groups: Problem solving versus behavioral skills training. *Research on Social Work Practice*, *4*(2), 172–191. http://dx.doi.org/10.1177/104973159400400204.

Maguin, E., Zucker, R. A., & Fitzgerald, H. E. (1994). The path to alcohol problems through conduct problems: A family-based approach to very early intervention with risk. *Journal of Research on Adolescence*, *4*(2), 249–269. http://dx.doi.org/10.1207/s15327795jra0402_5.

Malti, T., Ribeaud, D., & Eisner, M. P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: A cluster randomized controlled trial. *Journal of Clinical Child and Adolescent Psychology*, *40*(5), 677–692. http://dx.doi.org/10.1080/15374416.2011.597084.

McCabe, K., Yeh, M., Lau, A., & Argote, C. B. (2012). Parent-child interaction therapy for Mexican Americans: Results of a pilot randomized clinical trial at follow-up. *Behavior Therapy*, *43*(3), 606–618. http://dx.doi.org/10.1016/j.beth.2011.11.001.

McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, *34*(4), 527–543. http://dx.doi.org/10.1007/s10802-006-9031-1.

Mejia, A., Calam, R., & Sanders, M. R. (2015). A pilot randomized controlled trial of a brief parenting intervention in low-resource settings in panama. *Prevention Science*, *16*(5), 707–717. http://dx.doi.org/10.1007/s11121-015-0551-1.

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. http://dx.doi.org/10.1016/j.cpr.2013.07.006.

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*, 364–386. http://dx.doi.org/10.1177/1094428106291059.

Mouton, B., & Roskam, I. (2015). Confident mothers, easier children: A quasi-experimental manipulation of mothers' self-efficacy. *Journal of Child and Family Studies*, *24*(8), 2485–2495. http://dx.doi.org/10.1007/s10826-014-0051-0.

Niccols, A. (2009). Immediate and short-term outcomes of the 'COPEing with toddler behaviour' parent group. *Journal of Child Psychology and Psychiatry*, *50*(5), 617–626. http://dx.doi.org/10.1111/j.1469-7610.2008.02007.x.

Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of Triple P-Positive Parenting Program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child and Family Psychology Review*, *11*, 114–144. http://dx.doi.org/10.1007/s10567-008-0033-0.

Patterson, G. R. (1982). *A social learning approach to family intervention, vol. 3: Coercive family process.* Eugene, OR: Castalia Pub. Co.

Perrin, E. C., Sheldrick, R. C., McMenamy, J. M., Henson, B. S., & Carter, A. S. (2014). Improving parenting skills for families of young children in pediatric settings: A randomized clinical trial. *JAMA Pediatrics*, *168*(1), 16–24. http://dx.doi.org/10.1001/jamapediatrics.2013.2919.

Piquero, A. R., Jennings, W. G., Diamond, B., Farrington, D. P., Tremblay, R. E., Welsh, B. C., & Reingle Gonzalez, J. M. (2016). A meta-analysis update on the effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology*. http://dx.doi.org/10.1007/s11292-016-9256-0 Advance online publication.

Reedtz, C., Handegård, B. H., & Mørch, W. T. (2011). Promoting positive parenting practices in primary pare: Outcomes and mechanisms of change in a randomized controlled risk reduction trial. *Scandinavian Journal of Psychology*, *52*(2), 131–137. http://dx.doi.org/10.1111/j.1467-9450.2010.00854.x.

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child Psychology*, *9*, 22–29. http://dx.doi.org/10.1080/15374418009532938.

Rothman, A. J. (2000). Toward a theory-based analysis of behavioral maintenance. *Health Psychology*, *19*(1), 64–69. http://dx.doi.org/10.1037/0278-6133.19.Suppl1.64.

Rushton, A., Monck, E., Leese, M., McCrone, P., & Sharac, J. (2010). Enhancing adoptive parenting: A randomized controlled trial. Clinical Child Psychology &amp. *Psychiatry*, *15*(4), 529–542. http://dx.doi.org/10.1177/1359104510365041.

Sanders, M. R. (1999). Triple P – Positive parenting program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, *2*(2), 71–90. http://dx.doi.org/10.1023/A:1021843613840.

Sandler, I., Schoenfelder, E., Wolchik, S., & MacKinnon, D. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology*, *62*, 239–-299. http://dx.doi.org/10.1146/annurev.psych. 121208.131619.

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: Nuffield Foundation.

Sheeber, L. B., & Johnson, J. H. (1994). Evaluation of a temperament-focused, parent-training program. *Journal of Clinical Child Psychology*, *23*(3), 249–259. http://dx.doi.org/10.1207/s15374424jccp2303_3.

Sigmarsdóttir, M., Degarmo, D. S., Forgatch, M. S., & Gumundsdóttir, E. V. (2013). Treatment effectiveness of PMTO for children's behavior problems in Iceland: Assessing parenting practices in a randomized controlled trial. *Scandinavian Journal of Psychology*, *54*(6), 468–476. http://dx.doi.org/10.1111/sjop.12078.

Sofronoff, K., Jahnel, D., & Sanders, M. (2011). Stepping stones Triple P seminars for parents of a child with a disability: A randomized controlled trial. *Research in Developmental Disabilities*, *32*, 2253–2262. http://dx.doi.org/10.1016/j.ridd.2011.07.046.

Somech, L. Y., & Elizur, Y. (2012). Promoting self-regulation and cooperation in pre-kindergarten children with conduct problems: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 412–422. http://dx.doi.org/10.1016/j.jaac.2012.01.019.

Sonuga-Barke, E. J., Daley, D., Thompson, M., Laver-Bradbury, C., & Weeks, A. (2001). Parent-based therapies for preschool attention-deficit/hyperactivity disorder: a randomized, controlled trial with a community sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(4), 402–408. http://dx.doi.org/10.1097/00004583-200104000-00008.

Sonuga-Barke, E. J., Brandeis, D., Cortese, S., Daley, D., Ferrin, M., Holtmann, M., et al. , & Dittmann, R. W.(2013). Nonpharmacological interventions for ADHD: systematic review and meta-analyses of randomized controlled trials of dietary and psychological treatments. *American Journal of Psychiatry*, *170*(3), 275–289. http://dx.doi.org/10.1176/appi.ajp.2012.12070991.

Spijkers, W., Jansen, D. E. M. C., & Reijneveld, S. A. (2013). Effectiveness of primary care. Triple P on child psychosocial problems in preventive child healthcare: a randomized controlled trial. *BMC Medicine*, *11*, 240. http://dx.doi.org/10.1186/1741-7015-11-240.

Stewart-Brown, S., Patterson, J., Mockford, C., Barlow, J., Klimes, I., & Pyper, C. (2004). Impact of a general practice based group parenting programme: Quantitative and qualitative results from a controlled trial at 12 months. *Archives of Disease in Childhood*, *89*(6), 519–525. http://dx.doi.org/10.1136/adc.2003.028365.

Thompson, M. J., Laver-Bradbury, C., Ayres, M., Le Poidevin, E., Mead, S., Dodds, C., ... Sonuga-Barke, E. J. (2009). A small-scale randomized controlled trial of the revised new forest parenting programme for preschoolers with attention deficit hyperactivity disorder. *European Child and Adolescent Psychiatry*, *18*(10), 605–616. http://dx.doi.org/10.1007/s00787-009-0020-0.

Trentacosta, C. J., Hyde, L. W., Shaw, D. S., Dishion, T. J., Gardner, F., & Wilson, M. (2008). The relations among cumulative risk, parenting, and behavior problems during early childhood. *Journal of Child Psychology and Psychiatry*, *49*(11), 1211–1219. http://dx.doi.org/10.1111/j.1469-7610.2008.01941.x.

van Aar, J., Asscher, J. J., Zijsltra, B. J. H., Deković, M., & Hoffenaar, P. J. (2015). Changes in parenting and child behavior after the home-start family support program: A 10 year follow-up. *Children and Youth Services Review*, *53*, 166–175. http://dx.doi.org/10.1016/j. childyouth.2015.03.029.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. http://dx.doi.org/10.18637/jss.v036.i03.

Webster-Stratton, C. (2006). *The incredible years. A trouble-shooting guide for parents of children aged 2–8 years.* Seattle: Incredible Years.

Weisz, J. R., & Kazdin, A. E. (2010). *Evidence-based psychotherapies for children and adolescents.* New York, NY: Guilford Press.

Westrupp, E. M., Northam, E., Lee, K. J., Scarth, S. E., & Cameron, F. (2015). Reducing and preventing internalizing and externalizing behavior problems in children with type 1 diabetes: A randomized controlled trial of the Triple P – positive parenting program. *Pediatric Diabetes*, *16*, 554–563. http://dx.doi.org/10.1111/pedi.12205.

Williamson, A. A., Knox, L., Guerra, N. G., & Williams, K. R. (2014). A pilot randomized trial of community-based parent training for immigrant Latina mothers. *American Journal of Community Psychology*, *53*(1–2), 47–59. http://dx.doi.org/10.1007/s10464-013-9612-4.