



# The Implementation Integrity of Parenting Programs: Which Aspects Are Most Important?

Fabrizia Giannotta<sup>1</sup> · Metin Özdemir<sup>2</sup> · Håkan Stattin<sup>3</sup>

Published online: 24 July 2019  
© The Author(s) 2019

## Abstract

**Background** The implementation of preventive interventions is considered a crucial aspect of their success. However, few studies have investigated which components of implementation are most important.

**Objective** We aimed to understand whether the components of implementation integrity—adherence, quality of delivery, dose, and participants’ involvement—influenced the effectiveness of four parenting programs. We also investigated factors associated with these components.

**Method** Data come from a national evaluation of parenting programs in Sweden. The study was a randomised controlled effectiveness trial, with a sample of 535 parents with 3–12-year-old children. Measures included parenting behaviors (angry outbursts, harsh parenting, attempts to understand, rewarding, and praising), child conduct problems (ECBI and SNAP-V), and measures tapping into the four components (adherence, quality of delivery, dose, and participant involvement).

**Results** We ran multilevel models and found that implementation quality (adherence and quality of delivery) did not influence the effects on parents and children. Conversely, participant involvement was associated with improvements in parenting and child conduct. Finally, parents’ perceptions of their leaders as supportive and understanding were associated with parents’ responsiveness and attendance.

**Conclusions** Our study highlights the importance of having actively engaged parents to maximise intervention effects.

**Keywords** Parenting programs · Implementation quality · Adherence · Quality of delivery · Dose

---

✉ Fabrizia Giannotta  
fabrizia.giannotta@gmail.com

<sup>1</sup> Division of Public Health Sciences, School of Health, Care and Social Welfare, Mälardalens University, Höskoleplan 1, 72123 Västerås, Sweden

<sup>2</sup> School of Law, Psychology and Social Work (JPS), Center for Developmental Research, Örebro University, Örebro, Sweden

<sup>3</sup> Department of Psychology, Uppsala University, Uppsala, Sweden

## Introduction

High-quality implementation of prevention programs has commonly been assumed to be a precondition of their effectiveness. The idea is that the recipients of a program can change and get better only if they are provided with exactly what the program promises. Nevertheless, very few studies have empirically examined this assumption. Existing studies are limited regarding evidence on which components of the implementation process are most important in ensuring the success of parenting programs delivered in ordinary service settings. In addition, the studies to date have focused on only one program at a time, providing insight into the role of implementation fidelity, but with limited generalizability across programs. Using data from an effectiveness trial encompassing four different programs, the present study attempts to examine which aspects of implementation integrity are associated with changes in parents' and children's behaviors. Also, the study attempts to elucidate the factors likely to be associated with implementation integrity.

Implementation integrity refers to “the degree to which treatment is delivered as intended” (Yeaton and Sechrest 1981). We adopted a definition of implementation integrity based on Dane and Schneider's (1998) conceptualization, according to which it has four components: (1) adherence or fidelity (the degree to which program components are delivered as prescribed); (2) dose (the frequency and quantity of program administration); (3) quality of delivery (the extent to which facilitators approach a theoretical ideal in transmitting the core components); and (4) participant involvement (the levels of participation and enthusiasm). Altogether, the four components ensure that a program is implemented as intended, which reduces what has been called Type-III error (Dobson and Cook 1980), which refers to the error of attributing the failure of a program to its components or theory, when it is due to defective implementation.

Parenting programs aim at improving parenting in order to ameliorate the relationship between parents and children and are widely recognized as effective in reducing children problems (for a review see Furlong et al. 2012). However, very few studies have investigated the effects of the components of implementation integrity on parenting program effectiveness. Among the few available, adherence to the program manual has been associated with intervention' effectiveness (for a review, see Durlak and DuPre 2008), but not always (Breitenstein et al. 2010). Quality of delivery and participant involvement have been associated with better outcomes in parents and children (Eames et al. 2009; Forgatch et al. 2005), while dose has shown contradictory results (Dane and Schneider 1998). Overall, there is evidence that the different aspects of implementation integrity are related to parenting program effects to varying extents.

There are some major limitations in the field. First, there are few studies that have investigated all the components of implementation integrity together. As Domitrovich and Greenberg (2000), Dusenbury et al. (2003), and Durlak and DuPre (2008) have pointed out in their reviews, few intervention studies adopt more than two components of implementation integrity (usually adherence and dose), and even fewer have linked these components to program effects. Berkel et al. (2011) have called for an integrative approach that accounts for the components in order to understand which are most important for program effectiveness. Thus, studies distinguishing and evaluating the effects of each aspect of implementation integrity on program outcomes are needed.

Moreover, the factors predicting high implementation integrity are not completely clear. Some authors have investigated the association between adherence and group leaders' training, concluding that the more precisely group leaders are trained, the more likely

they will implement a program with high fidelity (Rohrbach et al. 2010; Seng et al. 2006). Also, some characteristics of group leaders have been shown to be related to participants' attendance (dose). For instance, racial and socioeconomic similarity between participants and group leaders is associated with leaders' therapeutic engagement (Orrell-Valente et al. 1999), which in turn is related to higher rates of retention. Finally, although some studies have associated parental involvement with program adherence (Breitenstein et al. 2010), they have considered the influence of just one component at a time. A narrow approach limits knowledge of the relative predictive role of each factor, so a more comprehensive approach is required.

## The Current Study

The aims of this study are twofold. First, we aimed to understand whether the components of implementation integrity—adherence, quality of implementation, dose, and participant involvement—affect the effectiveness of parenting programs. To achieve this goal, we used the four most common programs in Sweden when the present evaluation started. Three of these programs are, to some extent, behaviorally based (Comet, Cope, Incredible Years), while one is non-behavioral (Connect). Finding effects of implementation integrity across different types of programs permits drawing conclusions that are not limited to a specific program but are applicable to parenting programs in general. By contrast with previous studies, we assessed the dimensions of implementation integrity at both group and individual level. Moreover, we adopted a multi-informant approach, combining observational data, team-leader reports, and parent-reports, rather than focusing solely on leader reports, which have been shown to overestimate implementation integrity (Dusenbury et al. 2005). Next, we examined the role of implementation integrity for the program outcomes using data from an effectiveness trial. Some researchers have argued that the programs found to be effective in efficacy trials may fail to function as well when they are delivered in regular service settings (effectiveness trials) due to poor implementation integrity (Bumbarger and Perkins 2008). Nevertheless, since most studies focusing on implementation integrity have been based on efficacy trials, we chose to test our research questions in an effectiveness trial. Second, we investigated factors thought to be related to good implementation integrity. In keeping with Berkel et al.'s propositions (2011), we examined both leader-related aspects (adherence, quality of delivery) and participant-related aspects (participant involvement, dose). The former are likely to be dependent on features of the facilitators, such as gender, age, education, and experience, whereas the latter may be primarily dependent on participants' perceptions of their facilitators and program.

## Method

### Design and Procedure

The present study is part of a larger project, *The National Comparison of Parenting Programs*, which aims to evaluate the effects on disruptive child behaviors of the most commonly used, manual-based parenting programs in Sweden. The study was designed as a randomized controlled effectiveness trial with pre- and post-test, one two-year follow up after post-test. Given the focus on implementation, the current paper is based on the measurements at pre- and post-test. The behavioral parenting programs considered were Cope

(Cunningham 2006), Incredible Years (Webster-Stratton et al. 2004), and Comet (Kling et al. 2006), a Swedish program similar to Patterson's Parent Management Training—Oregon Model, and one non-behavioral, attachment-based program, Connect (Moretti and Obsuth 2009), was also included (see Table 1). Parents were randomly assigned to a program or a control condition and they were unaware that different programs were available.

In order to reduce barriers to participation, in each administrative region, the programs were offered by the human services units (e.g., schools, social welfare agencies, and child and adolescent psychiatry clinics) to all the parents in need. Most parents had contacted a unit on their own, but a few were recruited through advertisements about the availability of parenting programs in their communities (which was also a part of normal routine in these communities). However, fewer parents started on the Incredible Years program (75.4%) than on the other programs. This was because of organizational problems as some of the parents recruited for Incredible Years had to travel long distances to take part in the program, and as a result, many chose not to attend. The procedures have been described in detail elsewhere (*blind for review*). A total of 104 parenting groups were run by 76 pairs of team leaders. Parents completed a questionnaire before and immediately after the intervention. After program completion, they were asked questions concerning their commitment, their satisfaction, and the competence of the group leaders.

## Participants

Parents of 749 children participated. The children's ages ranged from 3 to 12 years, with average age 7.70 years ( $SD=2.60$ ). They were randomly assigned to one of the four parenting programs or to a control condition (for the randomization procedures, see Stattin et al. 2015). Only parents who participated in one of the programs were included, thereby excluding the parents in a waitlist control condition or in a self-help condition (where parents read a book). For the present study, we used the report of one parent for each family. If both parents attended the meetings, we selected the parent who had participated in most sessions of the program as the primary reporter. If the number of attendances was equal between parents, we chose the mother. Overall, mothers were the primary reporters (85%). The final sample comprised 535 parents, with an average age of 37.7 years ( $SD=7.51$ ), ranging from age 20, to age 60. About three out of four were married or cohabiting (74%), and the rest were single parents. In most cases (89%), both parents were born in one of the Scandinavian countries. The average monthly household income after tax was 30,000–40,000 SEK (\$3500–\$4700). There were 6.1% whose monthly incomes were as low as 0–10,000 SEK (\$0–\$1200), and 24.9% had an income higher than 50,000 SEK (\$5900). Only the 6.3% of the parents acknowledged that their monthly income was not fully adequate. Finally, 45.5% of the parents had completed some university-level education, and 9% had only a compulsory-school education.

Parents attending the parenting programs did not differ with regard to marital status, monthly income, economic strain, or educational level. Because Connect was only provided for parents of children older than 9, parents participating in Connect were older and had older children than parents participating in the other programs (see Stattin et al. 2015).

One hundred and eleven team leaders, in 76 team-leader pairs, delivered the programs. All leaders received specific pre-project training. Their mean age was 49 ( $SD=8.5$ ), and 80% ( $N=94$ ) were women. The majority had a university degree (95%,  $N=106$ ), the rest a high-school diploma.

**Table 1** Description of the parenting programs' aims and format

Dimensions	Comet	Cope	Incredible years	Connect
Aims	To decrease negative child behaviors, including ADHD, ADD, and ODD	To decrease negative child behaviors, including ADHD, ADD, and ODD	To decrease negative child behaviors, including ODD	To decrease negative preteens' and teens' behavioral problems (CD, aggressiveness, violence, antisocial behavior, delinquency) and mental health issues (concurrent anxiety and depression, substance use problems)
Age range	3–12 years	3–12 years	3–8 years	9–16 years
Sessions	11 2.5-h weekly sessions Groups of 10–12 parents (6 families)	10 1-h weekly sessions Groups of maximum 25 parents	12 2.5-h weekly sessions Groups of 10–14 parents	10 1-h weekly sessions Groups of 12–14 parents

## Measures

### Outcomes of the Program

**Parenting Competence** The 17-item Parenting Sense of Competence Scale (PSOC, Johnston and Mash 1989) was used to assess competence in parenting. Higher scores indicate higher competence. Cronbach's alphas for subscales were .81 and .95 at T1 and T2, respectively.

**Parents' Reactions** Parents' reactions to child misbehavior were assessed on five scales: *Attempted to understand* (5-item), *Angry outbursts* (5 item) (Stattin et al. 2011), *Harsh parenting* (7-item), *Rewarding* (2-item), and *Praising* (2-item) (Webster-Stratton et al. 2001). Higher scores indicate higher frequency of parenting reactions. Cronbach's alphas were .69 for *Attempted understanding*, .79 for *Angry outbursts*, and .63 for *Harsh parenting* at T1, and .68, .76, and .72 at T2, respectively. Correlations between the two items measuring *Praising* were .64 ( $p < .001$ ) at T1, and .58 ( $p < .001$ ) at T2, while correlations between the two items measuring *Rewarding* were .69 ( $p < .001$ ) at T1, and .64 ( $p < .001$ ) at T2.

**Children's Externalizing Problems** Eyberg's Child Behavior Inventory (ECBI) (Eyberg and Ross 1978), which comprises an *Intensity* and a *Problem Scale*, was used to assess children's externalizing problems. The *Intensity Scale* assesses the frequency of 36 externalizing behaviors, and the *Problem Scale* the extent to which parents consider each of the externalizing behaviors to be problematic. The alphas for the Intensity Scale were .93 at T1 and .94 at T2, and for the Problem Scale .91 on both occasions. The Swanson, Nolan and Pelham Rating Scale (SNAP-IV) (Swanson et al. 1992) was used to assess *inattention*, *hyperactivity/impulsivity* and *oppositional defiant disorder (ODD)* (Kazdin et al. 1989). Cronbach's alpha was .91 for inattention, .92 for hyperactivity/impulsivity, and .91 for ODD at T1, and .92, .91, and .91, respectively, at T2. For all measures, higher scores indicate greater child problems.

### Dimensions of Implementation Integrity

*Participants' involvement* was assessed through parent reports after program completion. In keeping with Dane and Schneider's suggestion that involvement represents "levels of participation and enthusiasm", we used an item indicating parents' satisfaction with the program and an item assessing the quantity of homework the parents completed at home (see Table 2 for a description). These two items were analyzed separately because they represent two different aspects of participants' involvement, as it was confirmed by their moderate correlation ( $r = .40$ ).

*Dose* was assessed using the records of attendance kept by group leaders (see Table 2). Because there were variations in the number of sessions for each program, we converted attendance rates into an ordinal scale referring to the percentage of sessions attended by parents (1 = less than 25%, 2 = 26–50%, 3 = 51–75%, 4 = more than 75%).

*Adherence and quality of delivery* were assessed through observations made by independent raters following the definition of Dane and Schneider (1998). Three sessions per group were randomly selected and video-recorded, resulting in 228 videotaped group sessions. Of these, 56 (25%) were randomly extracted, stratified by program, and coded by

**Table 2** Measures of implementation integrity

Dimensions	Measures (source of information)	Questions/items	Answers	M (SD)
Dose	Attendance (group leaders' records)	How many times did you participate in the program?	1 = less than 25% 2 = 26–50% 3 = 51–75% 4 = more than 75%	3.40 (.81)
Participants' responsiveness	Satisfaction with the program (parent report)	What do you think about the parental education you attended?	1 = very bad 3 = bad 4 = neither good or bad 5 = good 6 = very good	4.46 (.74)
	Homework (parent report)	How many exercises did you do at home?	1—I did no exercises 3—I did 1–2 exercises 4—I did 3–4 exercises 5—I did more than 4 exercises	3.94 (1.26)
Adherence and quality of delivery	Quality of the implementation (Observations)	To what extent did the group leaders follow the manual? How would you rate the quality of the group leaders' work? Think here primarily about how well the group leaders managed to convey the theoretical foundations of the program To what extent was a clear agenda followed? To what extent was there a clear description of the objectives of the sessions? To what extent did group leaders show enthusiasm?	10—points Likert scale: 1—very bad 10—very good	7.65 (1.44)

independent experts with extensive experiences of being a group leader and trainer of other leaders. To train the expert raters the following procedure was adopted. First, two raters for each program rated five videotaped sessions together until they approached consensus (these five sessions were not part of the tapes that were finally rated). Then, they independently rated about half of the sessions. Next, to avoid drift, the experts together rated another five videotapes to maintain consensus in their ratings. Finally, they independently rated the rest of the videotapes. We used averaged scores across raters. The items used to rate the video-recorded sessions are shown in Table 2.

*Adherence* (i.e., the extent to which the group leader followed the program manual) was measured with one item, while *quality of the delivery* was assessed with 4 items (see Table 2 for a description) that ranged from 1 (not at all) to 10 (totally). Interrater agreement, as indicated by the correlation between the ratings of the independent assessors, was high ( $r = .84$ ). The number of coded sessions for Comet was 17, for Cope 14, for Incredible Years 7, and for Connect 18. As the number of parents who started were 172 for Comet, 175 for Cope, 92 and 196 for Connect, the percentage of the coded sessions were about equal for each of the programs.

### Factors Related to Implementation Integrity

**Parents' Perceptions of Group Leaders** Parents rated their leaders at the end of the program. They reported the extent to which they could lead the group, support parents, and understand parents' problems, using one item for each behavior. Responses were rated on a 5-point scale ranging from 1, *not at all*, to 5, *fully*.

**Team Leaders' Characteristics** The team leaders were asked to state their gender, age, and level of education, and also asked whether they were specialized in a relevant area, such as psychotherapy. Because each group had two team leaders, we used gender composition (both females, both males, or mixed gender), average age, and average education, and aggregate specialization (i.e., none specialized, only one specialized, and both specialized) to represent the team-leader pairs.

### Statistical Analyses

First, we investigated whether *adherence* and *quality of delivery* represent two different dimensions. The dimensions were highly correlated, and a confirmatory factor analysis (CFA) showed that adherence and quality of delivery were parts of the same construct [ $\chi^2(4) = 3.62$ ,  $p > .05$ ; CFI = 1.00; RMSEA = .00; SRMR = .01]. Therefore, we combined these two dimensions into an *implementation quality* aggregate score by computing the mean of the ratings of the five items.

We used two-level multilevel regression models to address the first study question—how implementation fidelity is related to changes in parents' behaviors and competence, and children's behavior problems. The observations are nested in parenting groups and include the group-level implementation-quality measure. Clustering may lead to inflated Type-I error if not treated properly (Duncan et al. 2006). Thus, we used multilevel modeling with two levels in MPLus with the maximum likelihood robust (MLR) estimator (Muthén and Muthén 1998–2012): group level (Level 2) and individual level (Level 1). In all models, we controlled for pre-test levels of child and parent outcomes, type of program, and child and parents' age.



When an observation is missing at group level in nested data, the observations of the individual are also considered missing for the cluster in question. Therefore, we imputed missing data at group level using all available data external to the study models using a multiple-imputation technique (Enders 2010). Implementation quality and team leader characteristics were group level data. Because implementation quality was assessed by the ratings of a subset of video recordings, data were available for 58% of the groups. Overall, 70% of the participants were attending these groups. Thus, 70% of the individual level observations had also valid group level data on implementation quality. The main source of missing data for individual level observations was longitudinal attrition. The rate of longitudinal attrition was between 14 and 18%. We imputed five data sets, and merged them with the individual-level data. To examine the predictors of the dimensions of implementation fidelity, we fitted linear regression models using the TYPE = COMPLEX option in MPlus and the MLR estimator (Muthén and Muthén 1998–2012). The TYPE = COMPLEX option provides corrected standard-error estimates, reducing potential bias in test statistics due to clustering. In these models, we entered dummy-coded variables to control for differences across the programs.

## Results

### Descriptive Analyses

All programs were implemented with relatively high quality, with the lowest mean rating, on a 10-point scale, of  $M=7.03$  for Incredible Years. Despite the high quality of implementation, there were some differences between the programs. Cope and Comet had the highest quality, while Incredible Years had the lowest (see Table 3). Parents in Connect, followed by Comet, showed less absenteeism than parents in Cope and Incredible Years. Also, parents attending Comet completed their homework more often than those attending the other programs. Finally, Comet parents were more satisfied than parents attending Cope, Connect and Incredible Years. In sum, implementation integrity was generally high for all the programs. However, Cope and Comet were implemented to a higher standard than the other programs, and Comet was most appreciated by parents.

Finally, we computed the correlations between attendance to program sessions and family structure (1 = married and cohabiting 0 = single parent), and between attendance to program sessions and child age. The correlations were  $r = .015$ , n.s., and  $r = -.009$ , n.s., respectively, suggesting that attendance to program was not associated to family structure and child age.

**Table 3** Comparisons between the parent-training programs on the dimensions of implementation integrity

	Comet	Connect	IY	Cope	$F(3, 573)$	$p$	$\eta^2$
Implementation quality	7.61 <sup>a</sup>	7.15 <sup>b</sup>	7.03 <sup>b</sup>	8.65 <sup>c</sup>	50.04	< .001	.21
Dose	3.60 <sup>a</sup>	3.80 <sup>b</sup>	3.44 <sup>a</sup>	3.24 <sup>c</sup>	15.21	< .001	.08
Completion of homework	4.69 <sup>a</sup>	3.40 <sup>b</sup>	4.39 <sup>b</sup>	4.19 <sup>c</sup>	80.15	< .001	.30
Parent satisfaction	4.72 <sup>a</sup>	4.36 <sup>b</sup>	4.59 <sup>b</sup>	4.36 <sup>b</sup>	12.10	< .001	.06

Multiple F-test:  $F(12, 1716) = 36.81$ ,  $p < .001$ ,  $\eta^2 = .21$ . Different subscripts refer to significant differences between mean values, while same subscripts denote no significant difference

## Is Implementation Integrity Associated with Changes in Parenting and Child Problem Behaviors?

We examined the associations between the components of implementation integrity and changes in the parent and child outcomes in multilevel models. Implementation quality was entered as a group-level variable (Level 2), while parents' involvement and attendance were entered at individual level (Level 1).

### Effects of Implementation Integrity on Parent Outcomes

Implementation quality at group level and attendance at individual level were not significantly related to changes in parenting behaviors or parents' sense of competence (Table 4). By contrast, parents' involvement was significantly related to positive changes in parenting. Specifically, parents who completed their homework decreased most in angry outbursts ( $B = -.04, p < .05$ ), and increased in their use of praise ( $B = .14, p < .01$ ) and reward ( $B = .18, p < .01$ ), and in sense of parenting competence ( $B = .08, p < .05$ ). Finally, parents' satisfaction with the program significantly predicted decreases in harsh parenting ( $B = -.07, p < .01$ ), and increases in sense of parenting competence ( $B = .16, p < .01$ ). In sum, parents' involvement, i.e. satisfaction and homework completion, affected rates of change in parent behaviors and competence due to participation, whereas group-level implementation quality did not.

### Effects of Implementation Integrity on Child Outcomes

Implementation quality was not associated with changes in child problem behaviors and ADHD symptoms (see Table 5). But, the more parents were satisfied with their program, the more they reported reductions in their children's ECBI intensity ( $B = -.15, p < .001$ ) and problem ( $B = -.04, p < .01$ ) scores, inattention ( $B = -.10, p < .001$ ), and ODD symptoms ( $B = -.09, p < .05$ ). However, neither attendance nor homework completion predicted changes in child problem behaviors or ADHD symptoms. In sum, parents' satisfaction with their program predicted changes in child outcomes, whereas dose, homework completion, and group-level implementation quality did not predict program outcomes.

## Which are the Factors Associated with the Components of Implementation Integrity?

We examined predictors of the dimensions of implementation integrity. Specifically, we investigated whether leaders' characteristics (age, gender, education, specialization), and parents' perception of the leaders (leaders with good group management skills, supportive leaders, leaders that understand their problems) were predictors of dose, homework completion, and satisfaction with the program (Table 6). To account for differences across the programs, we entered dummy-coded variables into the models as controls. Thus, the unique effect of each predictor variable refers to its impact beyond differences due to the programs.

After controlling for differences across the programs, implementation quality seemed to be higher when team leaders had specialized training relevant to prevention ( $\beta = .30, p < .05$ ), and lower when leaders were older ( $\beta = -.31, p < .001$ ) and when the leadership pair comprised two women ( $\beta = -.12, p < .05$ ) rather than being of mixed gender. Also,

**Table 4** Predicting changes in parental outcomes from implementation quality, dose, parental responsiveness (completion of homework and satisfaction): multilevel models

	Angry outbursts	Harsh parenting	Attempted under-standing	Praise	Reward	Parenting competence
<b>Level 2</b>						
Implementation quality	.01 (.05)	-.02 (.13)	-.01 (.01)	-.04 (.18)	-.08 (.11)	-.02 (.03)
Residual	.00 (.01)	.00 (.02)	.00 (.01)	.00 (.03)	.01 (.13)	.00 (.01)
<b>Level 1</b>						
Dose	-.01 (.02)	-.02 (.03)	-.01 (.02)	-.02 (.05)	-.13 (.09)	-.06 (.04)
Homework	-.04 (.02)*	-.01 (.03)	.02 (.02)	.14 (.05)**	.18 (.06)**	.08 (.04)*
Satisfaction with the program	-.04 (.03)	-.07 (.03)*	.03 (.02)	.06 (.06)	-.03 (.08)	.16 (.05)***
Residual	.11 (.01)	.16 (.01)	.06 (.01)	.52 (.03)	.74 (.13)	.25 (.02)

Values presented in the table are unstandardized regression estimates, and standard errors are in parenthesis. Pre-test measurements of the outcome variables were included in the regression models. In addition, child age, parent age, and dummy-coded program variables were entered into the equations as control variables  
 \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

**Table 5** Predicting changes in child outcomes from implementation quality, dose, parental responsiveness (completion of homework and satisfaction), Multilevel models

	ECBI	ECBI problem	Inattention	Hyperactivity	Oppositional defiance
<b>Level 2</b>					
Implementation quality	-.01 (.02)	.01 (.01)	-.01 (.02)	-.01 (.02)	.02 (.02)
Residual	.01 (.02)	.01 (.01)	.00 (.01)	.00 (.01)	.00 (.01)
<b>Level 1</b>					
Dose	-.05 (.05)	-.00 (.01)	.06 (.05)	.05 (.03)	.04 (.04)
Homework	-.06 (.04)	-.02 (.01)	-.01 (.02)	-.03 (.02)	-.01 (.03)
Satisfaction with the program	-.15 (.04)***	-.04 (.01)**	-.10 (.03)***	-.07 (.04)	-.09 (.04)*
Residual	.27 (.02)	.03 (.01)	.15 (.01)	.17 (.02)	.21 (.03)

Values presented in the table are unstandardized regression estimates, and standard errors are in parenthesis. Pre-test measurements of the outcome variables were included in the regression models. In addition, child age, parent age, and dummy-coded program variables were entered into the equations as control variables

$p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

**Table 6** Predictors of dose, homework completion, and satisfaction with the program

	Implementation quality	Dose	Homework	Satisfaction with program
Age of team leaders	-.31***	.03	.02	-.02
Gender of team leaders				
Both women	-.12**	-.07	.01	.03
Both men	-.06	.02	.03	.02
Educational background				
Both have university education	-.07	-.04	.02	.01
Have specialized training	.30***	.05	.01	.01
Parents' appreciation of the leaders				
Group management skills	.00	.07	.05	.34***
Understanding problems of parents	.10	.15**	.13**	.04
Supportive of parents	-.07	.10	.09	.30***
R-Sqr	.31***	.16***	.37***	.39***

Dummy-coded program variables were entered into the equations to control for differences across the programs. Values presented in the table are standardized regression coefficients

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$

dose was related to parents' perceptions of their leaders. Specifically, parents perceiving leaders as understanding of their problems ( $\beta = .15$ ,  $p < .01$ ) attended more. Homework completion was positively predicted by parents perceiving their group leaders as understanding their problems ( $\beta = .13$ ,  $p < .01$ ). Finally, parents' program satisfaction was predicted by having supportive team leaders ( $\beta = .30$ ,  $p < .001$ ), and leaders with good group management skills ( $\beta = .34$ ,  $p < .001$ ). In sum, parents' perceptions of leaders as competent, supportive, and sensitive to their problems seem to be promotive of implementation integrity.

## Discussion

The aim of this study was to examine whether parents in parenting programs benefited more when their programs were well implemented. Specifically, we investigated the effects of all aspects of implementation integrity, namely implementation quality (adherence and quality of delivery), participant involvement (homework and satisfaction), and dose (attendance). In general, we did not find a significant effect of group-level implementation quality, neither an effect of dose, i.e. attendance, but we found an effect of parents' involvement, i.e. satisfaction with the program and homework completion. Independent of the number of sessions they attended, the more participants were satisfied and practiced what they learned during the sessions, the more they positively changed their way of parenting. Thus, our study suggests that the key component for the success of a program is parents' involvement during the sessions rather than simple attendance.

While the lack of effects of dose has been confirmed in other studies (e.g. Dane and Schneider 1998), the lack of effect of implementation quality is quite unexpected. However, this finding should be interpreted with caution. Group-level implementation quality was rated very highly across all the programs, and there was low variability in these

ratings, which may have resulted in a non-significant effect of group-level implementation quality on how much parents and children changed. Therefore, we are cautious about stating that implementation quality does not matter. Further studies are needed to test the role of group-level implementation quality on program outcomes using data with greater variability.

Contrary to the findings related to implementation quality, it emerges clearly that participants' involvement, which consists of homework completion and satisfaction with the program, might influence how much parents and children benefit. Independent of the quality of implementation, the parents who actively committed to and were satisfied with their program displayed more changes, such as increased feelings of being competent in parenting, improved parenting strategies, and decreased child problem behaviors. It does not come as a surprise that active participation is associated with the benefits of participating in a parenting program. Scholars have widely demonstrated, through reviews and meta-analyses, that interactive delivery methods are one of the principle elements in effective prevention (Nation et al. 2003; Tobler et al. 2000). The underlying assumption is that interactive methods are effective because they favor active participation. However, the assumption that active participation in a parenting program is related to higher program effectiveness has rarely, if ever, been tested. This study contributes to the literature by demonstrating empirically that parents need to be actively involved, for example through keeping practicing at home the techniques they learned during the program, if they want to obtain the maximum benefit.

How do parents become actively involved in a program? In our study, it emerged that they were more likely to be satisfied, do their homework, and attend when they perceived their group leaders as supportive and understanding. These qualities are among the requirements for a group leader to build up a "therapeutic alliance" with parents (for a review, see Ackerman and Hilsenroth 2003). In both individual and family-therapy settings, therapeutic alliance is an important predictor of both attendance (e.g. Orrell-Valente et al. 1999) and improvements in clients (e.g. Hogue et al. 2006). Our results are in line with this, but it is not clear why and how in a group of parents, some develop these views on their group leaders while others do not. In the current study, socio-cultural characteristics of the group leaders, such as sex and experience, did not seem to influence the association, as has been found in some other studies (e.g., Orrell-Valente et al. 1999). Future studies should investigate further the reasons why some parents perceive their leaders as supportive and others do not.

We also found some predictors of implementation quality. Our study suggests that mixed gender pairs of team leaders, and those with a specialization, such as therapist training, implement programs better than female pairs, and team leaders without a specialization. Leaders' age was negatively associated with quality of implementation. This result is in contrast with a recent study showing that older leaders are more likely to understand reasons for not changing the content of a program than younger and inexperienced leaders (Hill et al. 2007). However, in this study, only attitudes toward implementation were assessed. In our trial, in the majority of cases, group leaders delivered their programs in the manner to which they were accustomed. Consequently, older group leaders might have received training several years ago, by contrast with younger group leaders. For that reason, younger leaders might have been more sensitive to the importance of delivery of the programs without deviations from the program manual than the older leaders. However, this hypothesis cannot be confirmed in our study, and should be tested in future studies.

This study has some limitations. The first is related to the timing of assessments. Participants' involvement was assessed by parent reports at post-test. Parents were also

asked to report on their and their children's behaviors. It is possible that perceptions of changes from pre- and post-test would have affected their satisfaction with, attendance of, and commitment to the program, rather than the opposite. In other words, it is equally possible that the parents of children who benefited were more likely to be satisfied with, keep attending, and be actively involved in the programs. It is not possible to test the direction of effects in the current study. An ideal design to assess directionality would encompass measurements about parents and leaders following each program session. Nevertheless, such a measurement-intensive design would be difficult to implement in an effectiveness trial. Future studies may overcome this difficulty by using automated feedback technologies, with smart phones or tablet computers.

Another limitation is that we investigated only a limited set of the factors that might explain implementation integrity. Durlak and DuPre (2008) point out that implementation might be affected by many factors at both macro level (community factors, organizational factors) and micro level (characteristics of the innovation and the providers). We focused solely on micro-level factors, i.e., provider characteristics. Moreover, some of the measures, i.e. parents' satisfaction and homework completion, were single items and, as such, not ideal for the measurement of complex constructs, such as parents' involvement. Finally, because the participation was on a voluntary base and there were some organizational problems with the implementation of one of the programs (i.e. Incredible Years), generalizability of the results to all the parents (e.g. high and low-income families) cannot be guaranteed. However, we limited this problem by offering the programs to all the parents and contacting directly some of them, as previously described. Future studies should use better measures for micro-levels factors and account for the influences of macro-level factors on the different aspects of implementation integrity.

As well as limitations, this study has some strengths. It represents one of the first attempts to assess the impact of each component of program integrity simultaneously, which allowed us to understand the relative impact of each component. Moreover, the components were examined across different types of parenting programs, which makes our results likely to apply to parenting programs in general. Finally, it is one of the few studies that have assessed the role of implementation fidelity within an effectiveness trial, which allows us to draw conclusions that are applicable in real-life settings.

To conclude, our study highlights the importance of the active participation of parents in maximizing the positive effects of parenting programs. Group leaders with good training and empathic skills may be the key to promoting parents' involvement.

**Acknowledgements** Open access funding provided by Mälardarens University. This study was funded by The National Board of Health and Welfare (Socialstyrelsen, Sweden), n. 01-12,042/2008.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical Approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed Consent** Informed consent was obtained from all individual participants included in the study.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review, 23*(1), 1–33. [https://doi.org/10.1016/S0272-7358\(02\)00146-0](https://doi.org/10.1016/S0272-7358(02)00146-0).
- Berkel, C., Mauricio, A., Schoenfelder, E., & Sandler, I. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*(1), 23–33. <https://doi.org/10.1007/s11121-010-0186-1>.
- Breitenstein, S., Fogg, L., Garvey, C., Hill, C., Resnick, B., & Gross, D. (2010). Measuring implementation fidelity in a community-based parenting intervention. *Nursing Research, 59*(3), 158–165. <https://doi.org/10.1002/nur.20373>.
- Bumbarger, B., & Perkins, D. (2008). After randomised trials: Issues related to dissemination of evidence-based interventions. *Journal of Children's Services, 3*(2), 55–64. <https://doi.org/10.1108/17466660200800012>.
- Cunningham, C. (2006). Large group, community based, family-centered parent training. In R. A. Barkley & K. R. Murphy (Eds.), *Attention deficit hyperactivity disorder: A clinical workbook* (pp. 480–498). New York: Guilford Press.
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3).
- Dobson, D., & Cook, T. J. (1980). Avoiding type III error in program evaluation: Results from a field experiment. *Evaluation and Program Planning, 3*(4), 269–276. [https://doi.org/10.1016/0149-7189\(80\)90042-7](https://doi.org/10.1016/0149-7189(80)90042-7).
- Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*(2), 193–221. [https://doi.org/10.1207/s1532768xjepc1102\\_04](https://doi.org/10.1207/s1532768xjepc1102_04).
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and application*. Mahwah, NJ: Routledge.
- Durlak, J., & DuPre, E. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3–4), 327–350. <https://doi.org/10.1007/s10464-008-9165-0>.
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237–256. <https://doi.org/10.1093/her/18.2.237>.
- Dusenbury, L., Brannigan, R., Hansen, W. B., Walsh, J., & Falco, M. (2005). Quality of implementation: Developing measures crucial to understanding the diffusion of preventive interventions. *Health Education Research, 20*(3), 308–313. <https://doi.org/10.1093/her/cyg134>.
- Eames, C., Daley, D., Hutchings, J., Whitaker, C. J., Jones, K., Hughes, J. C., et al. (2009). Treatment fidelity as a predictor of behaviour change in parents attending group-based parent training. *Child: Care, Health and Development, 35*(5), 603–612. <https://doi.org/10.1111/j.1365-2214.2009.00975.x>.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.
- Eyberg, S. M., & Ross, A. W. (1978). Assessment of child-behavior problems—Validation of a new inventory. *Journal of Clinical Child Psychology, 7*(2), 113–116. <https://doi.org/10.1080/15374417809532835>.
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy, 36*(1), 3–13. [https://doi.org/10.1016/S0005-7894\(05\)80049-8](https://doi.org/10.1016/S0005-7894(05)80049-8).
- 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. <https://doi.org/10.1002/14651858.cd008225.pub2>.



- Hill, L., Maucione, K., & Hood, B. K. (2007). A focused approach to assessing program fidelity. *Prevention Science, 8*(1), 25–34. <https://doi.org/10.1007/s11121-006-0051-4>.
- Hogue, A., Dauber, S., Stambaugh, L. F., Cecero, J. J., & Liddle, H. A. (2006). Early therapeutic alliance and treatment outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology, 74*(1), 121–129. <https://doi.org/10.1037/0022-006X.74.1.121>.
- Johnston, C., & Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child and Adolescent Psychology, 18*(2), 167–175. [https://doi.org/10.1207/s15374424jccp1802\\_8](https://doi.org/10.1207/s15374424jccp1802_8).
- Kazdin, A. E., Bass, D., Siegel, T., & Thomas, C. (1989). Cognitive-behavioral therapy and relationship therapy in the treatment of children referred for antisocial behavior. *Journal of Consulting and Clinical Psychology, 57*(4), 522–535. <https://doi.org/10.1037/0022-006x.57.4.522>.
- Kling, A., Sundell, K., Melin, L., & Forster, M. (2006). Komet för föräldrar. *En randomiserad effektutvärdering av ett föräldraprogram för barns beteendeproblem. FoU-rapport, 14*.
- Moretti, M., & Obsuth, I. (2009). Effectiveness of an attachment-focused manualized intervention for parents of teens at risk for aggressive behaviour: The Connect Program. *Journal of Adolescence, 32*(6), 1347–1357. <https://doi.org/10.1016/j.adolescence.2009.07.013>.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., et al. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist, 58*(6–7), 449. <https://doi.org/10.1037/0003-066X.58.6-7.449>.
- Orrell-Valente, Jr, Laird, R. D., Bierman, K. L., Coie, J. D., & Pinderhughes, E. E. (1999). If it's offered, will they come? Influences on parents' participation in a community-based conduct problems prevention program. *American Journal of Community Psychology, 27*(6), 753–783. <https://doi.org/10.1023/a:1022258525075>.
- Rohrbach, L., Gunning, M., Sun, P., & Sussman, S. (2010). The project towards no drug abuse (TND) dissemination trial: Implementation fidelity and immediate outcomes. *Prevention Science, 11*(1), 77–88. <https://doi.org/10.1007/s11121-009-0151-z>.
- Seng, A. C., Prinz, R. J., & Sanders, M. R. (2006). The role of training variables in effective dissemination of evidence-based parenting interventions. *International Journal of Mental Health Promotion, 8*(4), 20–28. <https://doi.org/10.1080/14623730.2006.9721748>.
- Stattin, H., Enebrink, P., Özdemir, M., & Giannotta, F. (2015). A national evaluation of parenting programs in Sweden: The short-term effects using an RCT effectiveness design. *Journal of Consulting and Clinical Psychology, 83*(6), 1069–1084. <https://doi.org/10.1037/a0039328>.
- Stattin, H., Persson, S., Burk, W. J., & Kerr, M. (2011). Adolescents' perceptions of the democratic functioning in their families. *European Psychologist, 16*(1), 32. <https://doi.org/10.1027/1016-9040/a000039>.
- Swanson, J., Nolan, W., & Pelham, W. (1992). The SNAP-IV rating scale. <http://www.adhd.net>. November 14, 2009.
- Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *The Journal of Primary Prevention, 20*(4), 275–336. <https://doi.org/10.1023/A:1021314704811>.
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in Head Start. *Journal of Clinical Child Psychology, 30*(3), 283–302. [https://doi.org/10.1207/S15374424JCCP3003\\_2](https://doi.org/10.1207/S15374424JCCP3003_2).
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33*(1), 105–124. [https://doi.org/10.1207/S15374424JCCP3301\\_11](https://doi.org/10.1207/S15374424JCCP3301_11).
- Yeaton, W. H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. *Journal of Consulting and Clinical Psychology, 49*(2), 156–167. <https://doi.org/10.1037/0022-006x.49.2.156>.