



# Delivering a Parenting Program in South Africa: The Impact of Implementation on Outcomes

Yulia Shenderovich<sup>1</sup> · Manuel Eisner<sup>1</sup> · Lucie Cluver<sup>2,3</sup> · Jenny Doubt<sup>2</sup> · McKenzie Berezin<sup>4</sup> · Sybil Majokweni<sup>5</sup> · Aja Louise Murray<sup>1</sup>

Published online: 13 February 2019  
© The Author(s) 2019

## Abstract

**Objectives** Previous studies of parenting programs suggest that facilitator fidelity, participant attendance and engagement often influence treatment outcomes. While the number of parenting program evaluations has been growing in low- and middle-income countries, little is known about the implementation processes and their impact on participant outcomes in these settings.

**Methods** This study was nested within a cluster-randomised trial of a parenting program in South Africa. The paper aims to, first, describe the implementation of the intervention over 14 weeks. Second, using longitudinal multilevel analyses, the paper examines the impact of variation in observer-rated fidelity, attendance, and engagement on participant outcomes – parenting and maltreatment reported by caregivers and adolescents aged 10–18 ( $N = 270$  pairs), 14 outcome constructs.

**Results** Fidelity, attendance and participant engagement rates were similar to those reported in high-income country studies. However, the participation and implementation characteristics did not predict participant outcomes. This may be due to limited variation in dosage as home visits were comprehensively provided when participants could not attend group sessions, and fidelity was monitored by the implementers and researchers. One statistically significant predictor after the multiple testing correction was higher fidelity predicting an increase in adolescent-reported maltreatment at follow-up, possibly due to an increase in reporting (incidence rate ratio 1.33, 95% CI [1.19, 1.49],  $p < 0.01$ ).

**Conclusions** Our study confirms that a high quality of implementation can be achieved in a low-resource context. Suggestions for future research on parenting programs include examining therapeutic alliance alongside program fidelity and facilitator skill as well as systematically recording program adaptations.

**Keywords** Implementation · Process evaluation · Fidelity · Parenting · Child abuse

---

**Supplementary information** The online version of this article (<https://doi.org/10.1007/s10826-018-01319-y>) contains supplementary material, which is available to authorized users.

---

✉ Yulia Shenderovich  
y.shenderovich@gmail.com

<sup>1</sup> Institute of Criminology, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, UK

<sup>2</sup> Department of Social Policy and Intervention, Centre for Evidence-Based Intervention, University of Oxford, Oxford, UK

<sup>3</sup> Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa

<sup>4</sup> Department of Applied Psychology, New York University, New York, NY, USA

<sup>5</sup> Department of Psychiatry and Mental Health, HIV Mental Health Research Unit, University of Cape Town, Cape Town, South Africa

Parents' behavior has profound life-long effects on young people's health, wellbeing and education (Viner et al. 2012). Parenting interventions have been shown to improve child-caregiver relationships, promote positive parenting practices, and reduce harsh parenting as well as child maltreatment (e.g., see Barlow and Coren 2018). While most studies have been conducted in Western settings, a growing number of evaluations in low- and middle-income countries (LMIC) also suggest positive effects (Gardner et al. 2015; Knerr et al. 2013). Consequently, international agencies, such as the World Health Organization, have recommended parenting interventions among the key strategies for violence reduction within the family (World Health Organization 2016). Parenting support is seen as especially important for LMICs as most of the world's children live in LMICs and maltreatment tends to occur at higher rates than in high-income countries (Ward et al. 2016).

Converging evidence from interventions for a range of health and social outcomes suggests that interventions need to be delivered with sufficient quality and participant engagement to produce the intended results. In a widely cited review of prevention and health promotion programs for children (Durlak and DuPre 2008), authors found that in 45 out of 59 studies, there was a statistically significant positive relationship between implementation level and at least half of the intervention effects measured. A systematic review of group-based parenting programs for early-onset conduct disorders among children 3 to 12 years old found that studies reporting higher levels of implementation had larger intervention effects on measures, such as negative parenting practices, than the studies with lower implementation rates (Furlong et al. 2012). In line with these findings, the development and study of parenting interventions have been accompanied by an emphasis on the quality of intervention processes (Forgatch et al. 2005; Olofsson et al. 2016).

Berkel and co-authors have developed a helpful framework to guide research on the effects of implementation (Berkel et al. 2011), which focuses on two dimensions of implementation—facilitator and participant behavior—that are inter-related and can influence program outcomes. In assessing facilitator behavior, the authors identify three aspects: fidelity (intervention components delivered as prescribed), quality of delivery (facilitator teaching and process skills), and adaptation (facilitators modifying the intervention). Most of these terms have many alternative labels, for instance, fidelity is often referred to as adherence, and quality as facilitator competence, among other terms. To measure participant behavior (responsiveness), relevant aspects included in the model are attendance, the level of active engagement with the intervention, home practice and satisfaction.

Several parenting intervention studies have examined the effect of facilitator fidelity and quality of delivery on parenting outcomes, including studies of Incredible Years in the United Kingdom (Eames et al. 2009), Chicago Parent Program (Breitenstein et al. 2010) and Strengthening Families Program (Cantu et al. 2010) in the United States (US), Parent Management Training—Oregon Model in the US and Norway (Forgatch and DeGarmo 2011; Forgatch et al. 2005), Brief Parent Training in Norway (Kjøbli et al. 2012), and Growing Up Happily in the Family in Spain (Alvarez et al. 2016). Most, but not all of these studies (e.g., Breitenstein et al. 2010; Cantu et al. 2010) found higher facilitator fidelity or quality to be related to stronger improvements in some of the parenting or child behavior outcomes. A study of the Sinovuyo Kids project in South Africa focused on families with children 2–9 years old (Wessels 2017). In Sinovuyo Kids, the self-reported

performance of all facilitators was close to the maximum possible on the measure; thus, there was insufficient variation to examine the effect of fidelity on program outcomes.

The differences between the resources available in high-income countries (HICs) and LMICs may impact the extent to which interventions can be implemented to the desired standard. In HICs, facilitators in parenting and other psychosocial interventions often are expected to have graduate degrees and backgrounds in social work, psychology or education (Webster-Stratton and McCoy 2015). In LMICs, it is increasingly common for interventions to be delivered by lay workers due to the need for service scale-up and the limited numbers of available professional staff (Patel et al. 2011).

During parenting intervention sessions, participants can learn about, observe and practice various techniques, and each session has a specific topic. Therefore, if an intervention is efficacious, the participants who attend and actively engage with the intervention may be able to benefit the most. Indeed, studies of interventions, such as The Incredible Years (Baydar et al. 2003) and Chicago Parent Program in the US (Gross et al. 2009), found that more active attendance and engagement in sessions was linked to better intervention outcomes. In an evaluation of Fast Track in the US, engagement in sessions predicted several of the outcomes, while attendance did not, suggesting that attendance without active engagement may be insufficient to improve outcomes (Nix et al. 2009). However, not all studies found a statistically significant effect of participant attendance and engagement on treatment outcomes. Generally, even in studies that established a significant relation, the dose-response pattern was present for some, but not all outcomes (Alvarez et al. 2016; Gross et al. 2009; Wessels 2017; Weeland et al. 2017).

In summary, several previous parenting trials examined the role of implementation, attendance and engagement on outcomes, finding that the way the intervention was implemented and received was often significantly related to the variation in participant outcomes. Most of these trials focused on evaluating parenting programs aiming to improve child behavior. However, interventions targeting problem behavior in children are very similar in design to programs primarily addressing caregiver behavior to reduce child maltreatment (Knerr et al. 2013). The current study explores the implementation of Sinovuyo Teen, a parenting intervention for families with adolescents, within a randomised controlled trial (RCT) in South Africa. This paper first aims to describe the key dimensions of intervention implementation, and, second, to examine if the implementation measures in this study predict participant outcomes on child maltreatment and parenting behavior.

## Method

### Participants

This study was nested within an RCT in the Eastern Cape, South Africa. The Eastern Cape province is a historically disadvantaged part of South Africa, with poor infrastructure and high rates of poverty and unemployment. The trial was conducted between April 2015 and August 2016, and enrolled 552 families within 32 rural and 8 peri-urban study clusters, randomised into two arms. Families were identified through a range of sources, including self-referrals, local chieftains, community-selected representatives, schools, social services, as well as door-to-door visits. To be eligible for the study, families had to respond positively to one of the screening questions on whether there are conflicts between the caregiver and adolescent in the household, and to complete the two rounds of baseline assessments. The treatment clusters received the parenting program, described below. The control arm clusters received a hygiene information event. More details, including power calculations for the trial, are available in the study protocol (Cluver et al. 2016b).

### Procedure

The intervention manual was based on social learning principles, and developed drawing on existing research, consultations with experts and piloting in South Africa. Two pilot studies were conducted, including qualitative research with program facilitators and participants to incorporate their feedback in the revised manual and ensure program acceptability and relevance (Cluver et al. 2016a; Cluver et al. 2016c). Sessions were designed to be participatory and non-didactic, and they covered topics such as praise and relationship-building, managing emotions and solving family problems. The group sessions included both the adolescents and their caregivers to facilitate change in family relationships. The intervention was delivered by a local non-governmental organisation (NGO), and consisted of 14 weekly group sessions. In the ten joint sessions, caregivers and children worked together in the same room, while in the four separate sessions caregivers and adolescents were in parallel sessions, ideally out of each other's earshot, in order to promote an open discussion. In addition, the participants were given a home practice task after each session to practice new skills. For example, the facilitators asked adolescents and caregivers to give each other compliments during the week to practice the skill of offering specific labelled praise. For those who could not attend sessions, home visits were delivered by the facilitators with a short summary of the week's topic.

The facilitators were recruited by the implementing NGO to deliver the intervention during the RCT, and all

but two of the 25 facilitators had no previous experience with implementing a parenting intervention. Several facilitators were social workers seconded by the local government, and the others had various backgrounds in diverse fields, such as insurance, project management, arts and culture. Facilitators received a five-day training in collaborative facilitation methods and parenting principles, as well as on-going weekly day-long supervision and training on session content. Several strategies, such as session observations by the facilitators' trainer and NGO staff, were used to maintain treatment fidelity and quality.

### Measures

#### RCT data collection

Primary caregivers and adolescents completed self-report measures at baseline, post-test at one-month post-intervention (92% retention from baseline) and follow-up at 5–9 months post-intervention (97% retention from baseline). In addition, due to high population mobility in the study area, at the time of the follow-up data collection, 8% of the intervention arm adolescents who completed follow-up were neither living in the same house as the original caregiver nor spending most of the week together. Therefore these participants could not report on parenting in the original pair and were not included in the current analyses.

The randomised trial found improvements across several family and individual outcomes. Among the primary outcomes, caregivers reported reduced physical and emotional maltreatment, reduced use of corporal punishment and poor monitoring, as well as an increase in positive and involved parenting. Adolescents reported reduced maltreatment at post-test, but not follow-up, an increase in involved parenting, and a decrease in inconsistent discipline. Adolescents reported fewer intervention effects than their caregivers. Full details on the intervention effects are available in other publications (Cluver et al. 2018; Steinert et al. 2018), and basic descriptive information about the intervention arm is provided in Table 1.

#### Outcome measures

The outcome measures in this paper were the primary trial outcomes as specified in the study protocol. Adolescents' emotional and physical maltreatment, and neglect by caregivers were measured using a culturally-adapted version of the ISPCAN Child Abuse Screening Tool, ICAST-Trial (Meinck et al. 2018; Zolotor et al. 2009). At baseline, within the intervention arm, Cronbach alphas for the physical and emotional maltreatment scale was 0.78 (14 items) for caregiver report, and 0.87 (12 items) for adolescent report.

For neglect, the alphas were 0.62 (3 items) for caregiver report, and 0.77 (6 items) for adolescent report.

Parenting (poor parental monitoring, inconsistent discipline, corporal punishment, positive parenting and positive involved parenting) was measured using the Alabama Parenting Questionnaire (APQ; parent and adolescent versions), widely used internationally and previously used in South Africa (Essau, Sasagawa, Frick 2006; Lachman et al. 2014). For caregiver report, Cronbach alphas were 0.78 for positive parenting (6 items), 0.77 for involved parenting (10 items), 0.72 for poor monitoring (10 items), 0.74 for corporal punishment (3 items) and 0.55 for inconsistent discipline (6 items). For adolescent report, Cronbach alphas were 0.89 for positive parenting (6 items), 0.87 for involved parenting (10 items), 0.76 for poor monitoring (10 items), 0.68 (3 items) for corporal punishment and 0.68 for inconsistent discipline (6 items).

### Predictors

Implementation measures were collected by the Research Assistants (RAs) observing sessions in order to reduce the burden on the facilitators and participants and to avoid social desirability bias in self-reported implementation measures (Dusenbury et al. 2005). The RAs collected data

in 277 sessions, out of the total 279 sessions delivered, and 32% of the sessions were double-rated by two RAs. RA ratings were supplemented with comments that were reviewed during data analysis to validate and contextualize the quantitative trends. The information about home visits came from the records of the implementing NGO. We examine the following four implementation indicators as predictors of treatment outcomes.

*Group session attendance*—based on the program design, participation of both caregiver and adolescent in a family is necessary for intervention success. Therefore, we include the total number of sessions attended by adolescents and caregivers (the sum of the sessions each attended) as a candidate predictor of the changes in primary outcomes.

*Home visits*—as facilitators offered participants home visits after missed sessions, we also examined the effect of the total number of home visits delivered by facilitators to each household after missed sessions.

*Overall dosage*—we used the total sum of group sessions and home visits the family received to examine an overall dose-response effect. For simplicity, we treat group sessions at community venues and individual home sessions as equivalent. However, it is likely that the home sessions were not equivalent to the group sessions as the intervention relies on group interactions. In addition, the home sessions only lasted about 20 min each, about 5 times shorter than an average group session.

*Family engagement*—average of adolescent and caregiver engagement scores from all the sessions they attended. To measure the level of engagement in sessions, we used a behaviorally-anchored 3-point scale (1—Adolescent or caregiver was quiet or distracted most of the time; 2—Adolescent or caregiver participated in parts of the session; 3—Adolescent or caregiver participated through most of the session).

*Facilitator fidelity*—average rating of the facilitators in a cluster in sessions 1–14 (see Table 2). Slightly different items were used at the first and last sessions due to a different session structure. Cronbach alpha for the final scale

**Table 1** Baseline sample characteristics

| Variable  | Intervention arm   |
|---|--------------------|
| isiXhosa-speaking (African), <i>n</i> (%)                                   | 270 (100%)         |
| Rural village/Peri-urban township, <i>n</i> (%)                             | 224 (83%)/46 (17%) |
| Caregiver age, mean( <i>SD</i> )  | 49 years (15.2)    |
| Caregiver gender  | 262 (97%) female   |
| Child age, mean ( <i>SD</i> )   | 14 years (2.3)     |
| Child gender  | 118 (44%) female   |
| No one working in the household, <i>n</i> (%)                               | 178 (66%)          |
| At least 2 days in the past week with not enough food at home, <i>n</i> (%) | 181 (67%)          |

**Table 2** Items used to assess session implementation

| Dimension of implementation | Item (scored 0–3)   |
|-----------------------------|---|
| Fidelity (adherence)        | Facilitators introduce core lessons at the beginning of session   |
|                             | Facilitators remind participants of core lessons where relevant during session  |
|                             | Facilitators involve participants in remembering core lessons at the end of session                                     |
|                             | Role play (positive and negative) (e.g. “Facilitators or participants perform a positive and a negative role play”)     |
| Quality (competence)        | Role play discussion (e.g. “Using the role play, facilitators demonstrate the importance of managing stress and anger”) |
|                             | Do the facilitators ask participants open-ended questions?  |
|                             | Are the facilitators judgemental towards the participants? (reverse-coded)  |

was 0.92 (6 items). These items were developed based on the program manual and in consultation with the program developers. The observation tools assessed fidelity by measuring how well, according to RA observation, the facilitators implemented the core activities in a session, such as introducing, reinforcing and summarizing core lessons, and performing and discussing role plays. Moreover, we assessed facilitator skill by measuring the extent to which facilitators encouraged an open and supportive discussion.

While in seven of the 20 intervention groups the original pair of facilitators delivered the entire program, in others, the same pair of facilitators did not always facilitate in the same groups. Given this variation, we focus on the average ratings of all the facilitators in a group, rather than looking at the impact of individual facilitators. To create the final scale measuring facilitator performance, we used exploratory factor analysis, which suggested a one-factor solution (Kjøbli et al. 2012). The item on whether the facilitator was judgemental did not correlate well with the other items, perhaps because facilitators showing critical judgements of participant behaviors were quite rare; hence, this item was not included in the final scale. As the final scale primarily measured the completion of intervention activities, we treat it as a measure of fidelity. The approach to measuring fidelity was focused on the function of intervention activities (Hawe et al. 2004). In other words, the data collection emphasized assessing whether the purpose of the session activities was achieved rather than whether the activities were simply completed or not.

### Research assistant training

Fifteen RAs were involved in the implementation data collection after a five-day training in observational research and data collection forms. Furthermore, daily supervisions were conducted to ensure consistency in completing the forms. All RAs were from the local area and most of them had completed at least secondary education.

### Data Analyses

For the double-coded sessions, we examined inter-rater reliability of the observational measures. The intra-class correlation coefficients for fidelity items ranged between 0.7–0.9, and for participant engagement, 0.8–0.9. As high inter-rater reliability was achieved on these measures, we used averages of two observations for analyses.

Fixed-effect regression analyses and Kruskal-Wallis tests were used to check if the implementation characteristics differed significantly between groups. To examine changes in individual engagement over time, we used time (session

number) as a predictor of engagement in a multi-level model. To examine the impact of implementation predictors on participant outcomes, we adopted a longitudinal data analysis approach, as recommended for trials with repeated measures (Moerbeek and Teerenstra 2016). To detect the effect of implementation characteristics, we used the following model:

$$\begin{aligned} Parenting_{it} = & \beta_{00} + \beta_1(Predictor_i) + \beta_2(PT_{it}) + \beta_3(FU_{it}) + \\ & \beta_4(Predictor_i * PT_{it}) + \beta_5(Predictor_i * FU_{it}) + \\ & \alpha_1(Stratification_i) + u_{0i} + \varepsilon_{it}, \end{aligned} \quad (1)$$

where Level 1 = occasion ( $t$ ), Level 2 = individual ( $i$ ), PT is the immediate post-test and FU is the follow-up.  $Parenting_{it}$  is the estimated value for parenting at time  $t$  for person  $i$ ,  $Predictor_i$  is the candidate predictor that is a characteristic of the attendance, engagement or implementation for individual  $i$ . The interactions are the key parameters of interest used to assess the effect of predictors. These interactions  $\beta_4$  and  $\beta_5$  allowed us to examine the effect of the predictors on change over time in the intervention arm. The last two terms represent the time- and person-level residuals.

We estimated separate effects at post-test and follow-up. To avoid non-essential collinearity, all predictors were grand-mean centred. We used a linear link function for parenting outcomes and a negative binomial function for maltreatment, neglect and corporal punishment outcomes to account for the skewed distribution of the harsh parent behaviors. The linear models were estimated using maximum likelihood estimation, drawing on all available data. The outputs of the negative binomial models are presented as incidence rate ratios (IRR). We tested whether random slopes for time were indicated for these models (Barr 2013). However, they either did not substantively improve model fit or—in a small number of cases—produced convergence issues and so were not included in the final model. Although this was a cluster randomised trial, due to low ICC at cluster level (under 5%) and a design effect under 2.0, we did not include a separate level for clusters (Peugh 2010). For ICC values, see Table 3. To obtain conservative estimates, we used robust clustered standard errors to account for the cluster-level correlations.

We accounted for multiple hypothesis testing using the Benjamini–Hochberg procedure (Benjamini and Hochberg 1995), treating all tests as one family. For each outcome, we report both the unadjusted  $p$ -values as well as the corrected  $q$ -values. All analyses were implemented in Stata/SE 14.2, except the ICC inter-rater reliability calculations and the multiple hypothesis testing adjustment implemented in R 3.3.0 (package FSA v0.8.20).

**Table 3** Schedule of measurement of primary outcomes and their ICC values

| Outcome measure   | Baseline | Post-test | Follow-up | ICC at person level | ICC at cluster level |
|---|----------|-----------|-----------|---------------------|----------------------|
| Physical and emotional maltreatment (adolescent report) | X        | X         | X         | 0.29                | 0.02                 |
| Physical and emotional maltreatment (caregiver report)  | X        | X         | X         | 0.04                | 0.02                 |
| Neglect (adolescent report)                             | X        |           | X         | 0.30                | 0.00                 |
| Neglect (caregiver report)                              | X        |           | X         | 0.02                | 0.02                 |
| Corporal punishment (adolescent report)                 | X        |           | X         | 0.04                | 0.00                 |
| Corporal punishment (caregiver report)                  | X        |           | X         | 0.00                | 0.00                 |
| APQ positive parenting (adolescent report)              | X        | X         | X         | 0.22                | 0.02                 |
| APQ positive parenting (caregiver report)               | X        | X         | X         | 0.12                | 0.00                 |
| APQ involvement (adolescent report)                     | X        | X         | X         | 0.27                | 0.05                 |
| APQ involvement (caregiver report)                      | X        | X         | X         | 0.16                | 0.01                 |
| APQ poor monitoring (adolescent report)                 | X        | X         | X         | 0.20                | 0.03                 |
| APQ poor monitoring (caregiver report)                  | X        | X         | X         | 0.11                | 0.01                 |
| APQ inconsistent discipline (adolescent report)         | X        |           | X         | 0.04                | 0.02                 |
| APQ inconsistent discipline (caregiver report)          | X        |           | X         | 0.11                | 0.00                 |

## Results

As the summaries demonstrate, the families received 91% of the sessions in total either via group sessions or home visits (see Table 4). The average group session attendance, however, was only 58%. The highest number of group and home sessions received by a family was 31 instead of 28 because a few families received additional home visits, even though they attended the group sessions. Family engagement in group sessions was on average 74%, and average facilitator fidelity was 83% of the possible maximum.

### Attendance, Participation and Implementation Over Time

Neither family attendance and engagement, nor fidelity, displayed systematic change over time (see Table 5 for fidelity ratings by week). It appears that the level of fidelity was influenced by the specific circumstances and the content of the session in a particular week. Based on the RA observations, some of the common challenges with regard to fidelity were with the adoption of participatory approaches (e.g., facilitators reading role plays instead of acting them out with participants), and facilitators missing or misunderstanding parts of the content.

### Attendance, Participation and Implementation Across Clusters

Overall, there was a statistically significant difference between clusters in the average participant attendance, engagement and overall treatment dose. In other work, we examined predictors of attendance and engagement (Shenderovich et al. 2018).

### The Relationship of Implementation and Participant Behavior

There was a positive statistically significant correlation between average participant engagement and fidelity at cluster level,  $r(18) = 0.53$ ,  $p < 0.05$ . However, when examined separately for caregivers and adolescents, this effect was only statistically significant for average adolescent engagement,  $r(18) = 0.62$ ,  $p < 0.01$ . Although facilitators have reported challenges implementing the program when attendance was low, fidelity was not correlated with the average attendance rate in a particular group,  $r(18) = 0.02$ ,  $p > 0.05$  (see Table 6 for cluster-level correlations).

### Adaptations of the Intervention

The open-ended comments included in the RA observation forms suggest some adaptations of the Sinovuyo Teen program by facilitators during implementation. While adaptation was not measured systematically, we provide examples of the common changes made by facilitators, as reported by the RA observers. In some groups, facilitators had to make adaptations for low attendance when activities were designed to be practiced in family dyads but only one person from a dyad attended. On other occasions, instead of starting a new topic, facilitators continued with lessons from previous weeks when they felt something was not covered. Lack of physical space for separate sessions in some locations did not provide a quiet and confidential space separate for adolescents and caregivers. We also observed shorter duration of sessions than prescribed (1.8 h instead of 3), but session duration was not correlated with fidelity (see Table 6).

**Table 4** Descriptive statistics of the implementation characteristics

| Intervention characteristic | <i>N</i> | <i>M</i> ( <i>SD</i> ) | ICC by cluster | Maximum planned | Percentage of maximum |
|-----------------------------|----------|------------------------|----------------|-----------------|-----------------------|
| Group sessions              | 270      | 16.10 (7.86)           | 0.23           | 28              | 58%                   |
| Group and home sessions     | 270      | 25.42 (6.30)           | 0.02           | 28              | 91%                   |
| Family-level engagement     | 257      | 2.21 (0.38)            | 0.15           | 3               | 74%                   |
| Fidelity                    | 20       | 14.95 (1.29)           | N/A            | 18              | 83%                   |

The maximum fidelity score in sessions 1 and 14 was 12, and in sessions 2–13 it was 18, therefore the scores sessions 1 and 14 were scaled proportionately

**Table 5** Session fidelity ratings by week from all 20 intervention groups

| Session number | Session topic   | Total score, mean | Total score, <i>SD</i> | % of maximum |
|----------------|---|-------------------|------------------------|--------------|
| 1              | Introducing the program & defining participant goals            | 14.45             | 2.19                   | 80%          |
| 2              | Building a positive relationship through spending time together | 15.10             | 2.31                   | 84%          |
| 3              | Praising each other   | 16.05             | 2.04                   | 89%          |
| 4              | Talking about emotions  | 14.26             | 3.95                   | 79%          |
| 5              | What do we do when we're angry?                                 | 14.90             | 2.26                   | 83%          |
| 6              | Problem solving: Putting out the fire                           | 14.76             | 3.56                   | 82%          |
| 7              | Motivation to save and making a budget with our money           | 13.79             | 4.10                   | 77%          |
| 8              | Dealing with problems without conflict I                        | 14.75             | 2.55                   | 82%          |
| 9              | Dealing with problems without conflict II                       | 15.17             | 1.97                   | 84%          |
| 10             | Establishing rules and routines                                 | 15.21             | 2.50                   | 85%          |
| 11             | Ways to save money & making a family saving plan                | 17.05             | 0.88                   | 95%          |
| 12             | Keeping safe in the community                                   | 14.47             | 1.90                   | 80%          |
| 13             | Responding to crisis  | 15.55             | 1.83                   | 86%          |
| 14             | Widening Circles of Support                                     | 16.04             | 1.53                   | 89%          |

## Implementation Characteristics as Predictors of Outcomes

At both post-test and follow-up, there was no overall trend of the implementation characteristics affecting the outcomes in the intervention arm (see Tables 7 and 8). However, there were several statistically significant effects of implementation characteristics. A higher number of home visits was related to slightly larger reduction in maltreatment at post-test (caregiver report), while at follow-up, a higher total dose of the intervention was related to a larger reduction in inconsistent parenting (adolescent report) as well as a larger reduction in neglect scores (caregiver report). There was a statistically significant effect of family engagement, with higher family engagement predicting lower adolescent-reported involved parenting at post-test ( $p = 0.040$ ). Higher fidelity was marginally related to lower adolescent-reported positive parenting at post-test ( $p = 0.053$ ). Furthermore, groups with higher fidelity reported an increase in adolescent-reported

maltreatment and a smaller decrease in neglect, compared to the participants in groups with lower fidelity.

After adjusting for multiple hypothesis testing, the only remaining significant effect of implementation characteristics on participant outcomes was the effect of fidelity on adolescent-reported maltreatment. Contrary to our expectation, higher fidelity was linked to increased reports of maltreatment, IRR 1.33 95% CI [1.19, 1.49],  $p < 0.01$ .

## Discussion

In summary, the current study suggests that it was possible to deliver the Sinovuyo Teen intervention with a level of implementation comparable to studies in HICs, although it is not possible to compare precisely across studies due to different interventions and measures used. For example, in this study session fidelity was rated on average at 83%, participant attendance in group sessions at 58% and

**Table 6** Cluster-level correlations of implementation characteristics

|                 | Average caregiver engagement | Average adolescent engagement | Average family engagement | Average attendance | Average session length |
|-----------------|------------------------------|-------------------------------|---------------------------|--------------------|------------------------|
| Fidelity        | 0.21                         | 0.62**                        | 0.53*                     | 0.02               | 0.24                   |
| <i>p</i> -value | 0.347                        | 0.004                         | 0.020                     | 0.940              | 0.299                  |

\**p* < 0.05, \*\**p* < 0.01

**Table 7** The impact of implementation predictors on primary outcomes at post-test in the intervention arm

| Predictor                                     | Session attendance     | Home visits            | Sessions + home visits | Family engagement      | Cluster-level fidelity |
|---|------------------------|------------------------|------------------------|------------------------|------------------------|
| Outcome                                       | Coef. [95% CI]         |                        |                        |                        |                        |
| Maltreatment (adolescent report) <sup>a</sup> | IRR: 1.01 [0.98; 1.04] | IRR: 0.99 [0.96; 1.02] | IRR: 0.99 [0.94; 1.05] | IRR: 1.36 [0.77; 2.39] | IRR: 1.08 [0.94; 1.23] |
| <i>p</i> -value                               | 0.416                  | 0.363                  | 0.808                  | 0.292                  | 0.285                  |
| Adjusted <i>p</i> -value                      | 0.840                  | 0.822                  | 0.961                  | 0.822                  | 0.822                  |
| Maltreatment (caregiver report) <sup>a</sup>  | IRR: 0.99 [0.96; 1.03] | IRR: 1.00 [0.96; 1.05] | IRR: 0.99 [0.95; 1.03] | IRR: 1.09 [0.58; 2.00] | IRR: 0.94 [0.76; 1.17] |
| <i>p</i> -value                               | 0.672                  | 0.925                  | 0.692                  | 0.790                  | 0.590                  |
| Adjusted <i>p</i> -value                      | 0.915                  | 0.974                  | 0.915                  | 0.961                  | 0.908                  |
| Positive parenting (adolescent report)        | 0.00 [−0.11; 0.10]     | −0.06 [−0.20; 0.07]    | −0.06 [−0.18; 0.07]    | −1.84 [−3.82; 0.14]    | −0.66 [−1.32; 0.01]    |
| <i>p</i> -value                               | 0.939                  | 0.349                  | 0.368                  | 0.069                  | 0.053                  |
| Adjusted <i>p</i> -value                      | 0.974                  | 0.822                  | 0.822                  | 0.755                  | 0.755                  |
| Positive parenting (caregiver report)         | −0.01 [−0.13; 0.11]    | 0.04 [−0.04; 0.12]     | 0.06 [−0.12; 0.23]     | 0.42 [−1.88; 2.71]     | −0.12 [−0.69; 0.45]    |
| <i>p</i> -value                               | 0.879                  | 0.332                  | 0.508                  | 0.722                  | 0.676                  |
| Adjusted <i>p</i> -value                      | 0.974                  | 0.822                  | 0.865                  | 0.926                  | 0.915                  |
| Involved parenting (adolescent report)        | −0.05 [−0.18; 0.08]    | −0.10 [−0.29; 0.09]    | −0.19 [−0.46; 0.09]    | −2.14 [−4.18; −0.10]*  | −0.95 [−2.43; 0.54]    |
| <i>p</i> -value                               | 0.435                  | 0.312                  | 0.181                  | 0.040                  | 0.211                  |
| Adjusted <i>p</i> -value                      | 0.849                  | 0.822                  | 0.822                  | 0.755                  | 0.822                  |
| Involved parenting (caregiver report)         | 0.09 [−0.08; 0.25]     | −0.11 [−0.27; 0.05]    | 0.06 [−0.18; 0.30]     | 2.94 [−0.35; 6.23]     | 0.05 [−0.82; 0.92]     |
| <i>p</i> -value                               | 0.311                  | 0.176                  | 0.633                  | 0.080                  | 0.915                  |
| Adjusted <i>p</i> -value                      | 0.822                  | 0.822                  | 0.908                  | 0.755                  | 0.974                  |
| Poor parental monitoring (adolescent report)  | 0.10 [−0.08; 0.28]     | 0.04 [−0.12; 0.20]     | 0.23 [−0.08; 0.54]     | −1.43 [−4.43; 1.58]    | −0.14 [−1.31; 1.02]    |
| <i>p</i> -value                               | 0.278                  | 0.624                  | 0.141                  | 0.352                  | 0.810                  |
| Adjusted <i>p</i> -value                      | 0.822                  | 0.908                  | 0.822                  | 0.822                  | 0.961                  |
| Poor parental monitoring (caregiver report)   | −0.01 [−0.19; 0.17]    | 0.04 [−0.15; 0.23]     | 0.05 [−0.19; 0.28]     | −0.29 [−3.75; 3.18]    | −0.65 [−1.68; 0.38]    |
| <i>p</i> -value                               | 0.914                  | 0.697                  | 0.696                  | 0.871                  | 0.215                  |
| Adjusted <i>p</i> -value                      | 0.974                  | 0.915                  | 0.915                  | 0.974                  | 0.822                  |

<sup>a</sup>Negative binomial models, results presented as incidence rate ratios. All analyses controlled for rural-urban factor. \**p* < 0.05 for *p*-values before the correction for multiple hypotheses, †*p* < 0.05 for *p*-values after the correction

participant engagement at 74% of the maximum, whereas the Chicago Parent Program in the US reported average facilitator fidelity at 89%, group attendance at 50% and participant engagement at 82% (Breitenstein et al. 2010). This finding adds to the growing body of evidence from fields such as mental health that complex interventions can be delivered by lay staff with appropriate training and supervision (Kazdin and Rabbitt 2013).

Overall, the analyses did not identify the hypothesised pattern of impact of the intervention characteristics on participant outcomes. Different explanations may account for the absence of the expected effects. One explanation is that there was a very limited variation in the dosage for different participants within the RCT. As home visits were received nearly always when participants could not attend, the power to detect differences due to dosage was very limited. There was also no pattern of the effects of fidelity on the outcomes in the intervention arm. This may also be

due to low statistical power (Cantu et al. 2010). As intervention implementation within studies is usually of high quality and has limited variation, the statistical power is limited to detect the effects of fidelity, dose and engagement (Breitenstein et al. 2010; Nix et al. 2009).

After adjustment for multiple testing, there was one significant effect of treatment quality in the direction opposite to expected. Adolescents in clusters with higher quality of implementation reported an increase in maltreatment at follow-up, compared to adolescents in clusters with lower quality of implementation. One can only speculate as to the explanations for this finding. One possibility, given that there was no evidence of harm from the intervention in the quantitative and qualitative RCT data, is that the increased reporting of maltreatment by adolescents reflected their increased confidence and willingness to disclose their experiences at home. Indeed, as discussed above, average quality of implementation and adolescent

**Table 8** Implementation predictors at follow-up (intervention arm)

| Predictor<br>Outcome                                 | Session attendance     | Home visits             | Sessions and home visits | Family engagement      | Fidelity                            |
|--|------------------------|-------------------------|--------------------------|------------------------|-------------------------------------|
|  | Coef. [95% CI]         |                         |                          |                        |                                     |
| Maltreatment (adolescent report) <sup>a</sup>        | IRR: 0.98 [0.98; 1.04] | IRR: 1.03 [1.00; 1.05]  | IRR: 1.00 [0.96; 1.03]   | IRR: 1.28 [0.90; 1.81] | IRR: 1.33 [1.19; 1.49] <sup>†</sup> |
| <i>p</i> -value                                      | 0.240                  | 0.084                   | 0.817                    | 0.170                  | <0.001                              |
| Adjusted <i>p</i> -value                             | 0.822                  | 0.755                   | 0.961                    | 0.822                  | 0.004                               |
| Maltreatment (caregiver report) <sup>a</sup>         | IRR: 1.02 [0.98; 1.05] | IRR: 0.96 [0.92; 0.99]* | IRR: 0.98 [0.92; 1.04]   | IRR: 1.50 [0.84; 2.67] | IRR: 0.98 [0.79; 1.22]              |
| <i>p</i> -value                                      | 0.324                  | 0.016                   | 0.463                    | 0.167                  | 0.841                               |
| Adjusted <i>p</i> -value                             | 0.822                  | 0.436                   | 0.861                    | 0.822                  | 0.965                               |
| Neglect (adolescent report) <sup>a</sup>             | IRR: 0.99 [0.94; 1.03] | IRR: 1.01 [0.97; 1.06]  | IRR: 0.99 [0.95; 1.03]   | IRR: 0.67 [0.34; 1.32] | IRR: 1.54 [1.11; 2.14]*             |
| <i>p</i> -value                                      | 0.561                  | 0.630                   | 0.490                    | 0.249                  | 0.009                               |
| Adjusted <i>p</i> -value                             | 0.899                  | 0.908                   | 0.861                    | 0.822                  | 0.327                               |
| Neglect (caregiver report) <sup>a</sup>              | IRR: 0.91 [0.82; 1.00] | IRR: 1.00 [0.88; 1.14]  | IRR: 0.86 [0.77; 0.95]*  | IRR: 0.19 [0.03; 1.29] | IRR: 1.31 [0.67; 2.56]              |
| <i>p</i> -value                                      | 0.059                  | 0.956                   | 0.005                    | 0.090                  | 0.436                               |
| Adjusted <i>p</i> -value                             | 0.755                  | 0.974                   | 0.273                    | 0.755                  | 0.849                               |
| Corporal punishment (adolescent report) <sup>a</sup> | IRR: 0.99 [0.96; 1.03] | IRR: 1.00 [0.96; 1.04]  | IRR: 0.99 [0.93; 1.04]   | IRR: 1.15 [0.68; 1.94] | NA <sup>b</sup>                     |
| <i>p</i> -value                                      | 0.710                  | 0.927                   | 0.609                    | 0.609                  |                                     |
| Adjusted <i>p</i> -value                             | 0.921                  | 0.974                   | 0.908                    | 0.908                  |                                     |
| Corporal punishment (caregiver report) <sup>a</sup>  | IRR: 0.99 [0.95; 1.03] | IRR: 1.02 [0.97; 1.07]  | IRR: 1.00 [0.95; 1.05]   | IRR: 1.49 [0.86; 2.60] | IRR: 0.91 [0.75; 1.10]              |
| <i>p</i> -value                                      | 0.503                  | 0.953                   | 0.940                    | 0.155                  | 0.336                               |
| Adjusted <i>p</i> -value                             | 0.865                  | 0.974                   | 0.973                    | 0.822                  | 0.822                               |
| Positive parenting (adolescent report)               | 0.06 [−0.03; 0.16]     | −0.13 [−0.28; 0.01]     | −0.02 [−0.23; 0.18]      | −0.81 [−2.67; 1.04]    | −0.48 [−1.35; 0.40]                 |
| <i>p</i> -value                                      | 0.186                  | 0.077                   | 0.826                    | 0.390                  | 0.285                               |
| Adjusted <i>p</i> -value                             | 0.822                  | 0.755                   | 0.961                    | 0.822                  | 0.822                               |
| Positive parenting (caregiver report)                | −0.02 [−0.12; 0.09]    | −0.04 [−0.17; 0.09]     | −0.08 [−0.21; 0.04]      | 0.53 [−1.78; 2.85]     | 0.15 [−0.48; 0.78]                  |
| <i>p</i> -value                                      | 0.740                  | 0.545                   | 0.198                    | 0.652                  | 0.631                               |
| Adjusted <i>p</i> -value                             | 0.932                  | 0.898                   | 0.822                    | 0.915                  | 0.908                               |
| Involved parenting (adolescent report)               | 0.01 [−0.16; 0.18]     | −0.18 [−0.42; 0.07]     | −0.17 [−0.45; 0.10]      | −0.95 [−3.08; 1.18]    | −0.89 [−2.24; 0.46]                 |
| <i>p</i> -value                                      | 0.918                  | 0.155                   | 0.219                    | 0.382                  | 0.198                               |
| Adjusted <i>p</i> -value                             | 0.974                  | 0.822                   | 0.822                    | 0.822                  | 0.822                               |
| Involved parenting (caregiver report)                | 0.02 [−0.12; 0.16]     | −0.13 [−0.32; 0.07]     | −0.12 [−0.30; 0.06]      | 1.30 [−1.67; 4.26]     | 0.10 [−0.79; 0.98]                  |
| <i>p</i> -value                                      | 0.781                  | 0.204                   | 0.188                    | 0.392                  | 0.829                               |
| Adjusted <i>p</i> -value                             | 0.961                  | 0.822                   | 0.822                    | 0.822                  | 0.961                               |
| Poor parental monitoring (adolescent report)         | 0.03 [−0.14; 0.20]     | 0.09 [−0.16; 0.34]      | 0.15 [−0.10; 0.39]       | −1.35 [−5.44; 2.74]    | 0.16 [−0.64; 0.96]                  |
| <i>p</i> -value                                      | 0.744                  | 0.482                   | 0.239                    | 0.517                  | 0.692                               |
| Adjusted <i>p</i> -value                             | 0.932                  | 0.861                   | 0.822                    | 0.867                  | 0.915                               |
| Poor parental monitoring (caregiver report)          | −0.10 [−0.29; 0.09]    | 0.00 [−0.23; 0.23]      | −0.18 [−0.43; 0.08]      | −1.18 [−3.60; 1.24]    | −0.28 [−1.29; 0.72]                 |
| <i>p</i> -value                                      | 0.302                  | 0.999                   | 0.169                    | 0.338                  | 0.581                               |
| Adjusted <i>p</i> -value                             | 0.822                  | 0.999                   | 0.822                    | 0.822                  | 0.908                               |
| Inconsistent discipline (adolescent report)          | −0.07 [−0.19; 0.05]    | −0.05 [−0.19; 0.08]     | −0.18 [−0.35; 0.00]*     | −0.77 [−3.30; 1.76]    | 0.35 [−0.42; 1.12]                  |
| <i>p</i> -value                                      | 0.271                  | 0.459                   | 0.044                    | 0.552                  | 0.377                               |
| Adjusted <i>p</i> -value                             | 0.822                  | 0.861                   | 0.755                    | 0.898                  | 0.822                               |
| Inconsistent discipline (caregiver report)           | 0.00 [−0.08; 0.09]     | −0.04 [−0.17; 0.08]     | −0.04 [−0.15; 0.06]      | 0.49 [−0.90; 1.88]     | 0.02 [−0.53; 0.57]                  |
| <i>p</i> -value                                      | 0.977                  | 0.486                   | 0.408                    | 0.488                  | 0.938                               |
| Adjusted <i>p</i> -value                             | 0.986                  | 0.861                   | 0.839                    | 0.861                  | 0.974                               |

<sup>a</sup>Negative binomial models, results presented as incidence rate ratios, <sup>b</sup>model did not converge due to low prevalence. All analyses controlled for rural-urban factor. \* $p < 0.05$  for *p*-values before the correction for multiple hypotheses, <sup>†</sup> $p < 0.05$  for *p*-values after the correction

engagement in a cluster were strongly correlated. As highlighted in relation to intimate partner violence against women, violence is often under-reported and “willingness to disclose often improves with increased awareness about the definitions and extent of such abuse” (Kim et al. 2007: 1799). Similarly, in a recent study of the whole-school violence-reduction intervention the Good Schools Toolkit in Uganda, the researchers found an increase in girls’ experience of sexual violence as a result of the intervention, presumably due to increased reporting thanks to girls’ increased sense of safety to report violence (Devries et al. 2017). Although the intervention described in our study did

not explicitly address definitions of child maltreatment or children’s rights, it is possible that the intervention increased awareness of maltreatment by cultivating adolescents’ experience with non-violent problem-solving and communication. For example, one of the RA observers in a group session noted that adolescents present there “said they learned to tell their parents if they don’t like the way they [the caregivers] were treating them.”

Alternatively, another interpretation is that higher fidelity in some clusters may have meant less flexibility to meet the specific needs of the participants. Thus, perhaps the manual was followed more closely to the detriment of addressing

unexpected issues. Indeed, modifying interventions can be useful to make them more suitable for particular circumstances (Mowbray et al. 2003). In practice, adaptations always occur (Durlak 2015). Therefore, future parenting intervention research will benefit from planning how to systematically record and assess adaptations to evaluate their impact (Dusenbury et al. 2005). Furthermore, in complex interventions, it may not always be clear to facilitators how each activity or intervention process is expected to lead to a change in behavior. Providing clarifications for why various activities are included into the intervention manual and how they can be modified while preserving the core principles enables facilitators to adapt the intervention when necessary without undermining the ideas behind the program (Hill and Owens 2013).

The current evaluation also included a qualitative component exploring issues such as acceptability and mechanisms of change (Doubt et al. 2017). Intervention participants had primarily positive feedback, with most reporting improvements in their households, such as perceived reductions in violent discipline among caregivers and aggressive behaviors among children. Although high participant satisfaction was reported in interviews, attendance rates were lower than in some of the previous studies (e.g., Annan et al. 2017). Other work (Shenderovich et al. 2018) explores reasons for missed group sessions in Sinovuyo Teen. Overall, the study did not show evidence that the level of stressors, such as poverty, was related to attendance. There were several statistically significant predictors of attendance – for example, controlling for other variables, fewer sessions were attended in peri-urban clusters compared to rural areas, and by employed caregivers and male caregivers. Participants also reported other reasons for missed sessions, such as illness and community events. The offer of home visits following missed sessions may have also reduced the incentive to attend group meetings.

It is important to conduct replications by non-developers and research in circumstances that are typical of service settings (Forgatch and DeGarmo 2011). On the one hand, implementation tends to be closely supervised in a trial setting as compared to routine services. For example, in routine services interventions often experience program drift, meaning that implementation fidelity falls over time (Mowbray et al. 2003). On the other hand, interventions within trials may be delivered without the infrastructure available in routine services. For instance, in this trial, nearly all facilitators delivered the intervention for the first time during the RCT, immediately after the initial training and without any previous experience of working together.

There are also many important organisational factors in implementing any intervention that should be examined in future research on parenting interventions in LMICs. Factors such as facilitator recruitment, compensation,

incentives, and work conditions are likely to impact implementation (Fixsen et al. 2013). One model to promote fidelity at a larger scale that is used in parenting interventions is the certification of facilitators following training (Forgatch and DeGarmo 2011). Supervision and coaching of facilitators are likely to improve fidelity as well as help address facilitator burn-out and turnover (Sheidow et al. 2007). Various organisational models of delivering parenting programs can also be compared in experiments (Ashraf et al. 2014).

There is some evidence suggesting that monitoring implementation can boost its quality. Hence, implementation quality may be lower, if it is not monitored (Lachman et al. 2016). For instance, a review of 59 mentoring studies found that programs that monitored implementation had three times higher average effect sizes than programs without monitoring (DuBois et al. 2002). At the same time, this connection might not be causal; it is also possible that the projects that place an emphasis on quality have both higher overall quality of implementation and monitoring measures in place. Furthermore, as the Medical Research Council guidelines for process evaluations highlight, a key question for process evaluations is whether to communicate emerging findings as the intervention is implemented (Moore et al. 2015). In this study, as the intervention team was delivering the program for the first time, several emerging issues were communicated to the implementation team. Future process evaluations working in an effectiveness setting and with established providers are best suited to feed back all findings only at the end of the intervention (Hickey et al. 2016).

Even though only some of the parenting intervention studies report overall fidelity rates (Gardner et al. 2015), most of the published information on implementation of parenting interventions comes from experimental evaluations rather than routine delivery. However, as parenting programs are already widely implemented, it would be beneficial for future research to also draw on outcome monitoring in ordinary service delivery (Hurst et al. 2014) as this can greatly enlarge the pool of available data to examine the relationships between implementation and participant outcomes.

## Limitations

Our current analyses focus on quantitative implementation indicators, while providing context based on field notes of session observations, with qualitative research findings reported separately. As mentioned in the program description, participants who were unable to attend a session received a brief home visit from the facilitators with a summary of the week's content. Following missed sessions,

close to a third of the sessions were delivered through these individual visits, but it was not feasible to systematically observe home visits as observations in the home were considered too intrusive. It was also not possible to collect data on home practice completion, which is a program feature that parenting interventions rely on in order to reinforce the learning between sessions (Berkel et al. 2018; Ros et al. 2016). Furthermore, due to a limited number of intervention clusters (20), statistical power was limited, especially for fidelity as a cluster-level predictor, and it was not possible to explore predictors of fidelity. Future studies with larger samples would allow to test further hypotheses about mechanisms, such as whether effect of fidelity is mediated by participant engagement (Berkel et al. 2011).

Due to the relative brevity of the implementation measurement tool, our study did not examine facilitator fidelity and quality as two separate predictors, although they may have different effects on participant outcomes (Breitenstein et al. 2010; Humphrey et al. 2018). Future studies in LMICs could incorporate a more extensive subscale measuring facilitator competence. Another promising area is understanding therapeutic alliance – the interpersonal processes between the facilitator and participant—in the context of interventions delivered by lay staff in different cultural settings (e.g., see Elvins & Green, 2008; Schmidt et al. 2014).

This study also has some important strengths. It adds to the literature on the impact of implementation characteristics on outcomes in parenting interventions. Fidelity and engagement were observed by independent researchers rather than the implementers themselves, which provided a higher level of objectivity. In addition, data were collected from almost all intervention sessions rather than a limited sample. The fidelity, attendance, and engagement levels within an RCT may not be representative of all the replications in other settings by various agencies. However, these are helpful benchmarks that can be used for comparison and planning within the growing body of research examining parenting support in LMICs. Rigorously measuring implementation is a key step towards establishing evidence-informed parenting programs in LMICs.

**Acknowledgements** A number of research assistants contributed to the execution of this study, including Babalwa Mkati, Ntombizodwa Tsotso, Lungiswa Mzimeli, Kholiswa Mabizela, Veliswa Tokwe, Babalwa Mkati, Vuyolwethu Banzi, Deneo Sephulo, Mzuvukile Nocuza, Phelisa Mphimphilashe, Busisiwe Tusani, Zikhona Fihlani, Unathi Simandla, Ntombomzi Theodora, Noluvuyo Tsoko, Nonzukiso Bokwe, Nolitha Tshathu, Zikhona Capa, Mzoxolo Cannon, Pumla Myoyo, Simphiwe James Mahl, and Simphiwe Vimbayo. The authors also wish to thank Rocio Herrero Romero, Janina Steinert, Divane Nzima, Franziska Meinck, and the rest of the research and implementation teams for their dedicated work. Special thanks go to all the participants and their families.

**Author Contributions** Y.S., M.E., L.C., J.D. and M.B. contributed towards designing the study. Y.S., L.C., J.D., S.M. and M.B. contributed significantly towards data acquisition. Y.S. conducted data cleaning and analysis, and drafted the manuscript, and A.L.M. provided statistical advice. All authors participated in writing of the manuscript.

**Funding** Yulia Shenderovich was supported by the Cambridge Commonwealth, European & International Trust, research fellowship awarded by Global TIES for Children at New York University and funded by a grant from the Hewlett Foundation, Smuts Memorial Fund, managed by the University of Cambridge in memory of Jan Christiaan Smuts, and St John's College, Cambridge. The Sinovuyo Teen Study was funded by the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013, grant agreement 313421), UNICEF Innocenti Office of Research, the Leverhulme Trust (grant number PLP-2014-095), the University of Oxford's ESRC Impact Acceleration Account (K1311-KEA-004 & 1602-KEA-189), and the John Fell Fund (103/757).

## Compliance with Ethical Standards

**Conflict of Interest** L.C. and J.D. were involved in developing the Sinovuyo Caring Families Program for Parents and Teens, which is licensed under a Creative Commons 4.0 Non-commercial No Derivatives license. All other authors declare no financial relationships with any organisations that might have an interest in the submitted work and no other relationships or activities that could appear to have influenced the submitted work.

**Ethical Approval** The study was approved by the Research Ethics Committees at the University of Oxford, the University of Cape Town, and by the South African Departments of Social Development and Department of Basic Education.

**Informed Consent** Written informed consent was obtained from all individual participants included in the study. Both adolescents and caregivers needed to give informed consent for the dyad to participate.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**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 use, duplication, adaptation, distribution, and reproduction in any medium or format, as long as 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

- Alvarez, M., Rodrigo, M. J., & Byrne, S. (2016). What implementation components predict positive outcomes in a parenting program? *Research on Social Work Practice*, 28(2), 1–15. <https://doi.org/10.1177/1049731516640903>.
- Annan, J., Sim, A., Puffer, E. S., Salhi, C., & Betancourt, T. S. (2017). Improving mental health outcomes of Burmese migrant and displaced children in Thailand: a community-based randomized controlled trial of a parenting and family skills intervention. *Prevention Science*, 18(7), 793–803. <https://doi.org/10.1007/s11121-016-0728-2>.

- Ashraf, N., Bandiera, O., & Lee, S. S.-H. (2014). Awards unbundled: evidence from a natural field experiment. *Journal of Economic Behavior and Organization*, *100*, 44–63. <https://doi.org/10.1016/j.jebo.2014.01.001>.
- Barlow, J., & Coren, E. (2018). The effectiveness of parenting programs. *Research on Social Work Practice*, *28*(1), 99–102. <https://doi.org/10.1177/1049731517725184>.
- Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*, 328. <https://doi.org/10.3389/fpsyg.2013.00328>.
- Baydar, N., Reid, M. J., & Webster-Stratton, C. (2003). The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for head start mothers. *Child Development*, *74*(5), 1433–1453. <https://doi.org/10.1111/1467-8624.00616>.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289–300. <https://doi.org/10.2307/2346101>.
- Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science*, *12*(1), 23–33. <https://doi.org/10.1007/s11211-010-0186-1>.
- Berkel, C., Sandler, I. N., Wolchik, S. A., Brown, C. H., Gallo, C. G., Chiapa, A., & Jones, S. (2018). “Home Practice Is the Program”: Parents’ practice of program skills as predictors of outcomes in the new beginnings program effectiveness trial. *Prevention Science*, *19*(5), 663–673. <https://doi.org/10.1007/s11211-016-0738-0>.
- Breitenstein, S. M., Fogg, L., Garvey, C., Hill, C., Resnick, B., & Gross, D. (2010). Measuring implementation fidelity in a community-based parenting intervention. *Nursing Research*, *59*, 158–165. <https://doi.org/10.1097/NNR.0b013e3181dbb2e2>.
- Cantu, A., Hill, L., & Becker, L. (2010). Implementation quality of a family-focused preventive intervention in a community-based dissemination. *Journal of Children’s Services*, *5*(4), 18–30. <https://doi.org/10.5042/jcs.2010.0692>.
- Cluver, L. D., Lachman, J. M., Ward, C. L., Gardner, F., Peterson, T., Hutchings, J. M., & Redfern, A. A. (2016a). Development of a parenting support program to prevent abuse of adolescents in South Africa: findings from a pilot pre-post study. *Research on Social Work Practice*, *27*(7), 758–766. <https://doi.org/10.1177/1049731516628647>.
- Cluver, L. D., Meinck, F., Shenderovich, Y., Ward, C. L., Romero, R. H., Redfern, A., & Lechowicz, M. (2016b). A parenting programme to prevent abuse of adolescents in South Africa: Study protocol for a randomised controlled trial. *Trials*, *17*(328), 1–10. <https://doi.org/10.1186/s13063-016-1452-8>.
- Cluver, L. D., Meinck, F., Steinert, J., Shenderovich, Y., Doubt, J., Herrero-Romero, R., & Gardner, F. (2018). Parenting for lifelong health: a pragmatic cluster randomised controlled trial of a non-commercialised parenting programme for adolescents and their families in South Africa. *BMJ Global Health*, *3*(1), 1–16. <https://doi.org/10.1136/bmjgh-2017-000539>.
- Cluver, L. D., Meinck, F., Yakubovich, A., Doubt, J., Redfern, A., Ward, C., & Gardner, F. (2016c). Reducing child abuse amongst adolescents in low- and middle-income countries: a pre-post trial in South Africa. *BMC Public Health*, *16*(567), 1–11. <https://doi.org/10.1186/s12889-016-3262-z>.
- Devries, K. M., Knight, L., Allen, E., Parkes, J., Kyegombe, N., & Naker, D. (2017). Does the good schools toolkit reduce physical, sexual and emotional violence, and injuries, in girls and boys equally? A cluster-randomised controlled trial. *Prevention Science*, *18*, 839–853. <https://doi.org/10.1007/s11211-017-0775-3>.
- Doubt, J., Bray, R., Loening-Voysey, H., Cluver, L. D., Byrne, J., Nzima, D., & Medley, S. (2017). ‘It has changed’: understanding change in a parenting programme in South Africa. *Annals of Global Health*. <https://doi.org/10.1016/j.aogh.2017.10.021>.
- DuBois, D. L., Holloway, B. E., Valentine, J. C., & Cooper, H. (2002). Effectiveness of mentoring programs for youth: a meta-analytic review. *American Journal of Community Psychology*, *30*, 157–97. <http://www.ncbi.nlm.nih.gov/pubmed/12002242>
- Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science*, *16*, 1123–1127. <https://doi.org/10.1007/s11211-015-0606-3>.
- Durlak, J. A., & DuPre, E. P. (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*, 327–350. <https://doi.org/10.1007/s10464-008-9165-0>.
- 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*, 308–313. <https://doi.org/10.1093/her/cyg134>.
- Eames, C., Daley, D., Hutchings, J., Whitaker, C. J., Jones, K., Hughes, J. C., & Bywater, T. (2009). Treatment fidelity as a predictor of behaviour change in parents attending group-based parent training. *Child: Care, Health and Development*, *35*, 603–612. <https://doi.org/10.1111/j.1365-2214.2009.00975.x>.
- Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: an empirical review. *Clinical Psychology Review*, *28*, 1167–1187. <https://doi.org/10.1016/j.cpr.2008.04.002>.
- Fixsen, D. L., Blase, K. A., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children*, *79*, 213–230.
- Forgatch, M. S., & DeGarmo, D. S. (2011). Sustaining fidelity following the nationwide PMTO implementation in Norway. *Prevention Science*, *12*, 235–246. <https://doi.org/10.1007/s11211-011-0225-6>.
- 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).
- Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric properties of the Alabama parenting questionnaire. *Journal of Child and Family Studies*, *15*(5), 595–614.
- 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*. <https://doi.org/10.1002/ebch.1905>
- Gardner, F., Montgomery, P., & Knerr, W. (2015). Transporting evidence-based parenting programs for child problem behavior (age 3–10) between countries: systematic review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology*, *45*, 749–762. <https://doi.org/10.1080/15374416.2015.1015134>.
- Gross, D., Garvey, C., Julion, W., Fogg, L., Tucker, S., & Mokros, H. (2009). Efficacy of the Chicago Parent Program with low-income African American and Latino parents of young children. *Prevention Science*, *10*(1), 54–65. <https://doi.org/10.1007/s11211-008-0116-7>.
- Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how “out of control” can a randomised controlled trial be? *British Medical Journal*, *328*, 1561–1563. <https://doi.org/10.1136/bmj.328.7455.1561>.
- Hickey, G., McGilloway, S., Furlong, M., Leckey, Y., Bywater, T., & Donnelly, M. (2016). Understanding the implementation and effectiveness of a group-based early parenting intervention: a

- process evaluation protocol. *BMC Health Services Research*, 16 (490), 1–14. <https://doi.org/10.1186/s12913-016-1737-3>.
- Hill, L. G., & Owens, R. W. (2013). Component analysis of adherence in a family intervention. *Health Education*, 113(4), 264–280. <https://doi.org/10.1108/09654281311329222>.
- Humphrey, N., Barlow, A., & Lendrum, A. (2018). Quality matters: implementation moderates student outcomes in the PATHS curriculum. *Prevention Science*, 19(2), 197–208. <https://doi.org/10.1007/s11121-017-0802-4>.
- Hurst, A., Price, A., Walesby, R., Doolan, M., Lanham, W., & Ford, T. (2014). Routine outcome monitoring of evidence-based parenting programmes: Indications of effectiveness in a community context. *Journal of Children's Services*, 9(1), 58–74. <https://doi.org/10.1108/JCS-09-2013-0030>.
- Kazdin, A. E., & Rabbitt, S. M. (2013). Novel models for delivering mental health services and reducing the burdens of mental illness. *Clinical Psychological Science*, 1(2), 170–191. <https://doi.org/10.1177/2167702612463566>.
- Kim, J. C., Watts, C. H., Hargreaves, J. R., Ndhlovu, L. X., Phetla, G., Morison, L. A., & Pronyk, P. (2007). Understanding the impact of a microfinance-based intervention on women's empowerment and the reduction of intimate partner violence in South Africa. *American Journal of Public Health*, 97(10), 1794–1802. <https://doi.org/10.2105/AJPH.2006.095521>.
- Kjøbli, J., Bjørknes, R., & Askeland, E. (2012). Adherence to brief parent training as a predictor of parent and child outcomes in real-world settings. *Journal of Children's Services*, 7(3), 165–177. <https://doi.org/10.1108/17466661211261352>.
- Knerr, W., Gardner, F., & Cluver, L. D. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: a systematic review. *Prevention Science*, 14(4), 352–363. <https://doi.org/10.1007/s11121-012-0314-1>.
- Lachman, J. M., Cluver, L. D., Boyes, M. E., Kuo, C., & Casale, M. (2014). Positive parenting for positive parents: HIV/AIDS, poverty, caregiver depression, child behavior, and parenting in South Africa. *AIDS Care*, 26(3), 304–313. <https://doi.org/10.1080/09540121.2013.825368>.
- Lachman, J. M., Kelly, J., Cluver, L. D., Ward, C. L., Hutchings, J., & Gardner, F. (2016). Process evaluation of a parenting program for low-income families in South Africa. *Research on Social Work Practice*, 28(2), 188–202.
- Meinck, F., Boyes, M., Cluver, L. D., Ward, C., Schmidt, P., DeStone, S., & Dunne, M. (2018). Adaptation and psychometric properties of the ISPCAN Child Abuse Screening Tool for use in trials (ICAST-Trial) among South African adolescents and their primary caregivers. *Child Abuse and Neglect*, 82, 45–58.
- Moerbeek, M., & Teerenstra, S. (2016). *Power analysis of trials with multilevel data*. Boca Raton, FL: CRC Press.
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., & Baird, J. (2015). Process evaluation of complex interventions: medical research council guidance. *British Medical Journal*, 350, h1258 <https://doi.org/10.1136/bmj.h1258>.
- Mowbray, C. T., Holter, M. C., Gregory, B., & Bybee, D. (2003). Fidelity criteria: development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315–340. <https://doi.org/10.1177/109821400302400303>.
- Nix, R. L., Bierman, K. L., & McMahon, R. J. (2009). How attendance and quality of participation affect treatment response to parent management training. *Journal of Consulting and Clinical Psychology*, 77(3), 429–438. <https://doi.org/10.1037/a0015028>.
- Olofsson, V., Skoog, T., & Tillfors, M. (2016). Implementing group based parenting programs: a narrative review. *Children and Youth Services Review*, 69, 67–81. <https://doi.org/10.1016/j.chilcyouth.2016.07.004>.
- Patel, V., Chowdhary, N., Rahman, A., & Verdeli, H. (2011). Improving access to psychological treatments: lessons from developing countries. *Behaviour Research and Therapy*, 49(9), 523–528. <https://doi.org/10.1016/j.brat.2011.06.012>.
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48(1), 85–112. <https://doi.org/10.1016/j.jsp.2009.09.002>.
- Ros, R., Hernandez, J., Graziano, P. A., & Bagner, D. M. (2016). Parent training for children with or at risk for developmental delay: the role of parental homework completion. *Behavior Therapy*, 47(1), 1–13. <https://doi.org/10.1016/j.beth.2015.08.004>.
- Schmidt, F., Chomycz, S., Houlding, C., Kruse, A., & Franks, J. (2014). The association between therapeutic alliance and treatment outcomes in a group triple P intervention. *Journal of Child and Family Studies*, 23(8), 1337–1350. <https://doi.org/10.1007/s10826-013-9792-4>.
- Sheidow, A. J., Schoenwald, S. K., Wagner, H. R., Allred, C. A., & Burns, B. J. (2007). Predictors of workforce turnover in a transported treatment program. *Administration and Policy in Mental Health and Mental Health Services Research*, 34(1), 45–56. <https://doi.org/10.1007/s10488-006-0061-3>.
- Shenderovich, Y., Eisner, M., Cluver, L., Doubt, J., Berezin, M., Majokweni, S., & Murray, A. L. (2018). What affects attendance and engagement in a parenting program in South Africa? *Prevention Science*, 19(7), 977–986. <https://doi.org/10.1007/s11121-018-0941-2>.
- Steinert, J. I., Cluver, L. D., Meinck, F., Doubt, J., & Vollmer, S. (2018). Household economic strengthening through financial and psychosocial programming: evidence from a field experiment in South Africa. *Journal of Development Economics*, 134, 443–466. <https://doi.org/10.1016/j.jdeveco.2018.06.016>.
- Viner, R. M., Ozer, E. M., Denny, S., Marmot, M., Resnick, M., Fatusi, A., & Currie, C. (2012). Adolescence and the social determinants of health. *The Lancet*, 379(9826), 1641–1652. [https://doi.org/10.1016/s0140-6736\(12\)60149-4](https://doi.org/10.1016/s0140-6736(12)60149-4).
- Ward, C., Sanders, M. R., Gardner, F., Mikton, C. R., & Dawes, A. (2016). Preventing child maltreatment in low- and middle-income countries. Parent support programs have the potential to buffer the effects of poverty. *Child Abuse and Neglect*, 54, 97–107. <https://doi.org/10.1016/j.chiabu.2015.11.002>.
- Webster-Stratton, C., & McCoy, K. P. (2015). Bringing The Incredible Years programs to scale. *New Directions for Child and Adolescent Development*, 149, 81–95. <https://doi.org/10.1002/cad.20115>.
- Weeland, J., Chhangur, R. R., van der Giessen, D., Matthys, W., de Castro, B. O., & Overbeek, G. (2017). Intervention effectiveness of The Incredible Years: New insights into sociodemographic and intervention-based moderators. *Behavior Therapy*, 48(1), 1–18. <https://doi.org/10.1016/j.beth.2016.08.002>.
- Wessels, I. M. (2017). *Parental Engagement in Parent Training Interventions: Findings from the Sinovuyo Caring Families Project*. Cape Town, South Africa: University of Cape Town.
- World Health Organization. (2016). *INSPIRE: seven strategies for ending violence against children*. Luxembourg: World Health Organization.
- Zolotor, A. J., Runyan, D. K., Dunne, M. P., Jain, D., Péturs, H. R., Ramirez, C., & Isaeva, O. (2009). ISPCAN Child Abuse Screening Tool Children's Version (ICAST-C): Instrument development and multi-national pilot testing. *Child Abuse and Neglect*, 33, 833–841. <https://doi.org/10.1016/j.chiabu.2009.09.004>.