CHAPTER 7

FRIENDS FOR LIFE

IMPLEMENTATION OF AN INDICATED PREVENTION PROGRAM TARGETING CHILDHOOD ANXIETY AND DEPRESSION IN A NATURALISTIC SETTING

Mia P Kösters, Mai JM Chinapaw, Marieke Zwaanswijk, Marcel F van der Wal, Elisabeth MWJ Utens, Hans M Koot

Submitted for publication
OBJECTIVES When a prevention program is implemented in a naturalistic setting, deviations from protocol regularly occur, which may influence the program's effectiveness. We assessed the implementation of FRIENDS for Life, a school-based prevention program targeting childhood anxiety and depression, in a daily school practice, and examined the relation between implementation and program outcomes.

METHODS FRIENDS for Life was implemented as an indicated preventive intervention by mental health prevention workers in Amsterdam, the Netherlands. The sample consisted of 339 children with anxiety or depression symptoms, aged 8 to 13 years. We examined protocol adherence, quality of delivery, participant responsiveness, and exposure to the program using live observations. In addition, children's appraisal of the program was assessed.

RESULTS Prevention workers adhered not completely to the protocol. Children's participation in the program was good. We found few significant associations between program integrity and outcomes, and children rated the program more positively when protocol adherence was lower.

CONCLUSION Although prevention workers tend to adapt the program when implementing it in daily practice, this appears not to have negative effects on program outcomes. A highly protocolled intervention can be successfully transferred into daily school practice, even if it is adapted somewhat to practice.
INTRODUCTION

Anxiety and depression are common mental health problems in children (Beesdo, Knappe, & Pine, 2009). Symptoms of anxiety and depression in children and adolescents are associated with poor school performance, substance use and abuse, and suicidal behavior (Birmaher, Arbelaez, & Brent, 2002; Woodward & Fergusson, 2001). Moreover, children with untreated anxiety and depressive symptoms are at elevated risk for anxiety disorders and recurrent and more severe depressive episodes in later life (Beesdo et al., 2007; Birmaher et al., 2002; Fergusson & Woodward, 2002). Therefore, prevention of childhood anxiety and depression is important.

FRIENDS for Life is a program that aims to prevent anxiety and depression in children (Barrett, 2004a; Barrett, 2004b). The majority of FRIENDS for Life studies reported positive effects on anxiety or depression symptoms (e.g., Essau, Conradt, Sasagawa, & Ollendick, 2012; Barrett & Turner, 2001; Bernstein, Layne, Egan, & Tennison, 2005). However, much less is known about the implementation of FRIENDS for Life and the possible impact of implementation quality (program integrity) on the program’s effectiveness.

The way in which a program is implemented may influence its effectiveness in positive or negative ways (Dane & Schneider, 1998; Durlak & DuPre, 2008). Although the protocols of interventions – including FRIENDS for Life – generally thoroughly describe how the program should be implemented, deviations from protocols regularly occur when a program is executed outside the research setting. It is therefore important to evaluate to what extent program outcomes may be affected by program integrity.

Several studies investigating FRIENDS for Life as prevention program addressed program integrity. Most studies assessed adherence to protocol, and no study reported poor program integrity (e.g., Barrett, Lock, & Farrell, 2005; Essau et al., 2012; Rodgers & Dunsmuir, 2013). However, no study investigated the association between program integrity and effectiveness of FRIENDS for Life in an existing preventive setting. Furthermore, previous studies have several limitations that make it difficult to draw firm conclusions on program integrity and its influence on program outcomes.

Firstly, some studies used implementer-reported data about adherence to protocol (e.g., Essau et al., 2012; Barrett & Turner, 2001; Barrett et al., 2005). This kind of report may be prone to social desirable answers (Dane & Schneider, 1998). Secondly, not all studies quantified their results, but reported for instance that program integrity was high or that no deviations from the protocol were noted (Dadds, Spence, Holland, Barrett, & Laurens, 1997; Barrett, Moore, & Sonderegger, 2000; Lowry-Webster, Barrett, & Lock, 2003). However, quantification is needed to test the association between program integrity and program outcomes. Thirdly, the majority of studies of FRIENDS for Life only reported one or two aspects of program integrity, mainly adherence to protocol (e.g., Miller et al., 2011; Barrett & Turner, 2001). In literature it is recommended to investigate multiple aspects of program integrity (Durlak & DuPre, 2008; Dane & Schneider, 1998). For example, a program may be implemented completely according to protocol, but if participants were absent during numerous sessions, program integrity is still not optimal. Fourthly, up till now, the implementation of FRIENDS for
Life has only been studied in research-controlled settings, with extra training and evaluation for implementers (e.g., Barrett et al., 2005; Essau et al., 2012; Miller et al., 2011). Findings from these studies are not generalizable to implementation in naturalistic settings.

An additional aspect that may affect the implementation and effectiveness of a prevention program is participants’ appraisal of the program, also referred to as social validity. Even if an effective program is implemented with high program integrity, participants are likely to withdraw from the intervention if they do not like it. On the longer term, this may hamper sustainability and dissemination of the program. Previous studies showed that children and parents positively evaluated FRIENDS for Life (e.g., Barrett, Sonderegger, & Sonderegger, 2001; Cooley, Boyd, & Grados, 2004; Lowry-Webster et al., 2003). However, previous findings regarding the association between social validity and a reduction of symptoms of anxiety or depression are not univocal (Barrett et al., 2000; Essau, Conradt, & Ederer, 2004; Gallegos-Guajardo, Ruvalcaba-Romero, Garza-Tamez, & Villegas-Guinea, 2013).

The present study aims to address the abovementioned gaps in the literature with a comprehensive process evaluation of FRIENDS for Life as an indicated preventive school-based intervention, for children with elevated levels of anxiety or depression symptoms but not a clinical disorder. FRIENDS for Life has been implemented in Amsterdam, the Netherlands, as part of an existing prevention strategy since 2007. We included all FRIENDS for Life groups in two consecutive school years in a quasi-experimental trial, and asked the prevention workers to implement the program as they were used to do (Kösters et al., 2012). Prevention workers received no specific or additional training or supervision during the trial. In this way, we were able to study implementation and outcomes under naturalistic conditions. Results of the concurrent trial show that children who participated in FRIENDS for Life showed self-reported a strong reduction in anxiety and depression symptoms in comparison to controls, towards levels comparable to children from the general population at 12 months post-intervention (Kösters, Chinapaw, Zwaanswijk, van der Wal, & Koot, 2015).

We examined four aspects of program integrity using live observations: (a) adherence to protocol, (b) quality of delivery, (c) participant responsiveness (children’s participation in the sessions), and (d) exposure to the program (Dane & Schneider, 1998; Dusenbury, Brannigan, Falco, & Hansen, 2003). In addition, children’s appraisal of the program was assessed, as well as the association of each of these aspects with program outcomes. We aimed to address the following questions:

1. To what extent was the implementation of a highly protocolled prevention program affected by executing it in a naturalistic setting?
2. To what extent did implementation in a naturalistic setting affect program outcomes?
METHODS

This process evaluation is part of a larger quasi-experimental trial evaluating the effects of FRIENDS for Life (Kösters et al., 2012). In the present study, only data from the intervention groups were used. FRIENDS for Life was implemented in grade 6, 7 and 8 of elementary schools (comparable with grade 4, 5 and 6 in US schools) in Amsterdam, the Netherlands. During the school years 2010-2011 and 2011-2012, 35 FRIENDS for Life intervention groups started at 23 elementary schools in Amsterdam. Per school, up to 11 children with elevated symptoms of anxiety or depression were selected. More details about the selection procedure are described elsewhere (Kösters et al., 2012). Children and parents received information about the study and gave written permission if they wished to participate in the study.

The intervention group consisted of 339 children, of whom 6 did not start the program (main reason: second thoughts about participation) and 5 were excluded from FRIENDS by the prevention workers because of disruptive behavior. Participating children were 8-13 years old (M=10.6, SD=0.9), and 62% were girls. Children were of Dutch (20%), Turkish (12%), Moroccan (22%), Surinamese/Antillean (16%), other Western (8%), other non-Western (20%), and unknown (3%) descent.

INTERVENTION

FRIENDS for Life is based on cognitive behavior therapy (CBT) (Barrett, 2004a; Barrett, 2004b). Children learn how to cope with anxiety and depression by learning several skills and strategies. The program consists of 10 sessions, two booster sessions (one and three months after finishing the program), and two parent sessions.

In Amsterdam, the Netherlands, the Dutch translation of FRIENDS for Life (Utens & Ferdinand, 2006a; Utens & Ferdinand, 2006b) was implemented. Each group was led by two prevention workers (out of a pool of 21) from a local mental health organization. The 10 child sessions lasted 1.5 hours each and were conducted once a week during the school day. The implementation of booster and parent sessions deviated from the original protocol: as prevention workers noticed time constraints of schools and low attendance of parents, the implementation of only one booster session and one parent session has become common practice in Amsterdam over the years. The program started two times a year, after summer break and after Christmas break.

MEASURES

At the screening assessment (T1), children filled out the Revised Child Anxiety and Depression Scale (RCADS) (Chorpita, Yim, Moffitt, Umemoto, & Francis, 2000) and socio-demographic questions. At the end of the last (10th) FRIENDS for Life session (T2), children filled out the RCADS again and were asked to appraise the program.
Program integrity
Adherence to protocol, quality of delivery, and participant responsiveness were assessed by live observations. The first author, two trained Master’s level students and a trained research assistant conducted the observations. The first two to three sessions were observed by two observers (the first author and a student or assistant) to ensure that observations and scoring were done in a similar way. Differences were discussed and pilot observations were conducted until agreement was reached. These pilot observations were excluded from the present analyses. During the study period, a number of randomly selected sessions were again observed by two observers, to prevent coder drifting and to ensure that scoring continued in a similar way. The observations were not announced to the prevention workers. In total, 74 sessions (19% of the total number) were observed in 26 different FRIENDS for Life groups (2-5 observations per group), of which 14 sessions were observed by two persons.

Protocol adherence was assessed by the Program Integrity Checklist, which was specifically designed for FRIENDS for Life (Barrett, 1999). During the session, we scored the adherence to protocol components on a 4-point Likert scale (extremely well (3), moderately well (2), not very well (1), not at all (0)). Quality of delivery was assessed by scoring therapeutic skills with the Group Leader Integrity Checklist, also designed specifically for FRIENDS for Life, on a 4-point Likert scale (extremely well (3) - not at all (0)) (Barrett, 1999). The therapeutic skills included in the checklist were: positive reinforcement, specific feedback, self-disclosure, empathy, paraphrasing, summarization, and reflection. The scores of the two prevention workers were combined in one mean score per skill per session. To address participant responsiveness, we added an item to the observation checklist investigating the extent to which every child participated in the session, which was scored on a 4-point Likert scale (extremely well (3) - not at all (0)).

Exposure to the program was registered by the prevention workers. They logged the attendance of the children during program sessions and of parents during the parent sessions. In addition, they registered whether participating children had done their homework assignments. Unfortunately, parents’ attendance and homework were not well registered and could therefore not be analyzed.

Children’s appraisal of the program
Three items were presented to the children at T2. They rated on a 5-point Likert scale how useful (very little=0 to very much=4) and enjoyable (very boring=0 to a lot of fun=4) they considered the program. In addition, children were asked whether they would recommend the program to other children (yes/no). Answers were available for 229 of the 339 children since these questions were added during the course of the study. These 229 children did not differ from the total sample regarding gender, ethnicity, and T1 and T2 scores, but were slightly older ($M=10.7$ versus $M=10.5$, t-test $p=0.04$).
**Program outcomes**

Children completed the self-report RCADS at T1 and T2. This questionnaire consists of an anxiety scale (37 items) and a major depression scale (10 items) (Chorpita et al., 2000). Children indicate how often each item applies to them on a 4-point Likert scale (never, sometimes, often, always). In the present trial, Cronbach’s alphas were 0.95 at T1 and T2 for anxiety, and 0.78 at T1 and 0.83 at T2 for depression.

**Sociodemographic information**

Children were asked to fill in their date of birth, and their own and their parents’ country of birth. Ethnicity was based on the parents’ country of birth (cf. Statistics Netherlands, 2000).

**STATISTICAL ANALYSES**

To assess inter-rater reliability for observations, we calculated ICC agreement. This measure accounts for the degree of agreement (low, moderate, or complete) on a 4-point Likert scale and equals a weighted kappa (de Vet, Terwee, Mokkink, & Knol, 2011; Cicchetti, 1994). An ICC below 0.40 is defined as poor, between 0.40 and 0.59 as fair, between 0.60 and 0.74 as good, and between 0.75 and 1.00 as excellent (Cicchetti, 1994).

Program integrity measures and appraisal were analyzed as continuous variables (except for the question whether the child would recommend the program to others), as this has more statistical power and no standard cut-off values exist for low or high implementation (Durlak & DuPre, 2008; Dane & Schneider, 1998). Mean scores across sessions were calculated per group (adherence to protocol, quality of delivery) or child (participant responsiveness, exposure).

Firstly, associations between child characteristics and participant responsiveness, exposure and children’s appraisal of the program were examined using multilevel linear or logistic regression procedures. Characteristics were dichotomized: gender (boy=0, girl=1), age (8-10 years olds versus 11-13 years olds), ethnicity (Dutch versus non-Dutch), and severity of symptoms (lower three versus upper quartile) of the RCADS anxiety and depression scales. The quartiles of severity of initial symptoms were based on the RCADS scores of children included in the intervention and control group (n=496) (Kösters et al., 2015).

Secondly, associations between program integrity and program outcomes, and associations between children’s appraisal and program outcomes at post-intervention were examined, adjusted for baseline values (T1), using multilevel linear or logistic regression procedures.

Thirdly, associations between program integrity and children’s appraisal of the program were examined using multilevel linear regression procedures, adding adherence to protocol, quality of delivery, participant responsiveness, and exposure as independent variables to the model one by one.

ICCs, t-tests, and descriptive statistics were calculated using SPSS (IBM, version 21) and multilevel procedures were employed using MlwiN (Centre for Multilevel Modelling,
University of Bristol, United Kingdom, version 2.27). For multilevel procedures, we used a 3-level structure: individual, class, school; as children were nested within classes and classes were nested within schools.

RESULTS

INTER-RATER RELIABILITY

The inter-rater ICCs agreement for adherence to protocol were excellent (ranging from 0.81 to 0.90). The inter-rater ICCs for quality of delivery were good to excellent (ranging from 0.65 to 0.91); and the inter-rater ICCs for participant responsiveness were good (ranging from 0.61 to 0.68).

<table>
<thead>
<tr>
<th>VARIABLE (RANGE)</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adherence to protocol (0-3)</td>
<td>247</td>
<td>1.94</td>
<td>0.58</td>
</tr>
<tr>
<td>Quality of delivery (0-3)</td>
<td>247</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Positive reinforcement</td>
<td>247</td>
<td>2.87</td>
<td>0.25</td>
</tr>
<tr>
<td>- Specific feedback</td>
<td>247</td>
<td>2.48</td>
<td>0.54</td>
</tr>
<tr>
<td>- Self-disclosure</td>
<td>247</td>
<td>1.40</td>
<td>0.97</td>
</tr>
<tr>
<td>- Empathy</td>
<td>247</td>
<td>2.70</td>
<td>0.40</td>
</tr>
<tr>
<td>- Paraphrasing</td>
<td>247</td>
<td>1.44</td>
<td>0.80</td>
</tr>
<tr>
<td>- Summarization</td>
<td>247</td>
<td>0.38</td>
<td>0.69</td>
</tr>
<tr>
<td>- Reflection</td>
<td>247</td>
<td>1.06</td>
<td>1.00</td>
</tr>
<tr>
<td>Participant responsiveness (0-3)</td>
<td>247</td>
<td>2.51</td>
<td>0.43</td>
</tr>
<tr>
<td>Exposure</td>
<td>318</td>
<td>9.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Appraisal</td>
<td>229</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Usefulness (0-4)</td>
<td></td>
<td>3.11</td>
<td>0.83</td>
</tr>
<tr>
<td>- Enjoyableness (0-4)</td>
<td></td>
<td>3.44</td>
<td>0.70</td>
</tr>
<tr>
<td>- Recommend (yes)</td>
<td></td>
<td>91%</td>
<td></td>
</tr>
<tr>
<td>RCADS anxiety T1 (0-111)</td>
<td>332</td>
<td>38.9</td>
<td>23.0</td>
</tr>
<tr>
<td>RCADS anxiety T2</td>
<td>323</td>
<td>26.9</td>
<td>20.1</td>
</tr>
<tr>
<td>RCADS depression T1 (0-30)</td>
<td>336</td>
<td>8.7</td>
<td>5.2</td>
</tr>
<tr>
<td>RCADS depression T2</td>
<td>323</td>
<td>6.3</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Note: *M*=mean, *SD*=standard deviation. RCADS=Revised Child Anxiety and Depression Scale.
PROGRAM INTEGRITY

Table 1 reports means and standard deviations of adherence to protocol, quality of delivery, participant responsiveness, and exposure. Prevention workers adhered mainly but not completely to the protocol, and their quality of delivery varied considerably per skill. Children showed a high degree of responsiveness to the program and were present during most sessions.

Girls, $B=0.16$, 95% CI [0.06, 0.25], and children with higher initial levels of anxiety symptoms, $B=0.13$, 95% CI [0.02, 0.25] showed a higher degree of participation in the sessions. Age and ethnicity were not related with participant responsiveness. No child characteristics were associated with children’s attendance.

CHILDREN’S APPRAISAL OF FRIENDS FOR LIFE

Table 1 reports children’s appraisal of the program. Specific groups of children evaluated the usefulness of FRIENDS for Life differently. Boys found the program more useful than girls, $B=0.29$, 95% CI [0.07, 0.51], older children found it more useful than younger ones, $B=0.22$, 95% CI [0.01, 0.43], and non-Dutch children found it more useful than their Dutch peers $B=0.40$, 95% CI [0.15, 0.65]. For enjoyableness and the willingness to recommend the program to other children, we found no differences between specific subgroups of children.

PROGRAM INTEGRITY AND CHILDREN’S APPRAISAL IN RELATION TO PROGRAM OUTCOMES

We found few significant associations between program integrity and self-reported anxiety and depressive symptoms (Table 2). Only higher levels of prevention workers’ self-disclosure were associated with less self-reported anxiety at post-intervention, and higher levels of prevention workers’ specific feedback with less self-reported depression at post-intervention. Children who evaluated the program as useful had fewer anxiety symptoms post-intervention. No other significant associations were found between children’s appraisal and self-reported symptoms.

PROGRAM INTEGRITY AND CHILDREN’S APPRAISAL

Adherence to protocol was negatively related to children’s appraisal. The lower the overall adherence to protocol, the more positively children appraised program usefulness, $B=0.40$, 95% CI [0.14, 0.65] and enjoyableness, $B=0.31$, 95% CI [0.01, 0.60], and the more children were inclined to recommend the program to other children, $B=0.15$, 95% CI [0.02, 0.91]. Quality of delivery, participant responsiveness and exposure were not significantly associated with children’s appraisal of FRIENDS for Life.
### Table 2

**Relations between program integrity and program outcomes: results from multilevel linear regression analyses**

<table>
<thead>
<tr>
<th>OUTCOME</th>
<th>INTEGRITY INDICATOR</th>
<th>B</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RCADS Anxiety</strong></td>
<td>Adherence to protocol</td>
<td>-3.6</td>
<td>[-9.9, 2.7]</td>
<td>0.26</td>
</tr>
<tr>
<td>Quality of delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Positive reinforcement</td>
<td></td>
<td>-0.3</td>
<td>[-15.0, 14.5]</td>
<td>0.97</td>
</tr>
<tr>
<td>- Specific feedback</td>
<td></td>
<td>-5.4</td>
<td>[-12.2, 1.3]</td>
<td>0.11</td>
</tr>
<tr>
<td>- Self-disclosure</td>
<td></td>
<td>-3.6</td>
<td>[-6.9, -0.3]</td>
<td>0.03</td>
</tr>
<tr>
<td>- Empathy</td>
<td></td>
<td>4.8</td>
<td>[-3.1, 12.7]</td>
<td>0.23</td>
</tr>
<tr>
<td>- Paraphrasing</td>
<td></td>
<td>-0.7</td>
<td>[-5.7, 4.3]</td>
<td>0.78</td>
</tr>
<tr>
<td>- Summarization</td>
<td></td>
<td>0.2</td>
<td>[-4.9, 5.3]</td>
<td>0.93</td>
</tr>
<tr>
<td>- Reflection</td>
<td></td>
<td>0.3</td>
<td>[-2.4, 3.1]</td>
<td>0.80</td>
</tr>
<tr>
<td>Participant responsiveness</td>
<td></td>
<td>1.0</td>
<td>[-4.3, 6.2]</td>
<td>0.72</td>
</tr>
<tr>
<td>Exposure</td>
<td></td>
<td>-0.9</td>
<td>[-3.0, 1.2]</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>RCADS Depression</strong></td>
<td>Adherence to protocol</td>
<td>-1.1</td>
<td>[-2.4, 0.3]</td>
<td>0.13</td>
</tr>
<tr>
<td>Quality of delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Positive reinforcement</td>
<td></td>
<td>-3.6</td>
<td>[-7.3, 0.2]</td>
<td>0.06</td>
</tr>
<tr>
<td>- Specific feedback</td>
<td></td>
<td>-2.1</td>
<td>[-3.8, -0.3]</td>
<td>0.02</td>
</tr>
<tr>
<td>- Self-disclosure</td>
<td></td>
<td>-0.3</td>
<td>[-1.1, 0.6]</td>
<td>0.55</td>
</tr>
<tr>
<td>- Empathy</td>
<td></td>
<td>-0.4</td>
<td>[-2.4, 1.7]</td>
<td>0.72</td>
</tr>
<tr>
<td>- Paraphrasing</td>
<td></td>
<td>-0.7</td>
<td>[-2.1, 0.6]</td>
<td>0.27</td>
</tr>
<tr>
<td>- Summarization</td>
<td></td>
<td>-0.4</td>
<td>[-1.7, 0.9]</td>
<td>0.53</td>
</tr>
<tr>
<td>- Reflection</td>
<td></td>
<td>-0.1</td>
<td>[-0.8, 0.6]</td>
<td>0.79</td>
</tr>
<tr>
<td>Participant responsiveness</td>
<td></td>
<td>0.0</td>
<td>[-1.3, 1.3]</td>
<td>0.99</td>
</tr>
<tr>
<td>Exposure</td>
<td></td>
<td>-0.4</td>
<td>[-0.9, 0.2]</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>RCADS Anxiety</strong></td>
<td>Usefulness</td>
<td>-2.9</td>
<td>[-5.6, -0.2]</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Enjoyableness</td>
<td>2.7</td>
<td>[-0.7, 6.1]</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Recommend to other children</td>
<td></td>
<td>[-8.1, 8.0]</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>RCADS Depression</strong></td>
<td>Usefulness</td>
<td>-0.3</td>
<td>[-1.0, 0.4]</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Enjoyableness</td>
<td>0.4</td>
<td>[-0.5, 1.2]</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Recommend to other children</td>
<td></td>
<td>[-2.5, 1.5]</td>
<td>0.61</td>
</tr>
</tbody>
</table>
DISCUSSION

The present study investigated to what extent FRIENDS for Life, an indicated prevention program for childhood anxiety and depression, was implemented in a naturalistic setting, and to what extent the implementation affected program outcomes.

Prevention workers adhered to the protocol but did not execute the program exactly as prescribed. It is difficult to compare our results to previous studies, as these were based on different informants (implementers or participants), in different settings (controlled research setting), and with different samples (age range, sample size). Despite these differences, the majority of the previous results on adherence, as assessed by implementers or researchers, can be interpreted as more positive than our findings (e.g., Barrett, Sonderegger, & Xenos, 2003; Rodgers & Dunsmuir, 2013; Barrett et al., 2001). This may be explained by the fact that in the present study, FRIENDS for Life has been implemented for several years and was not investigated as part of a carefully controlled study, but as part of an existing prevention strategy. We asked prevention workers to implement the program as they had been doing for years.

Quality of delivery varied considerably between the investigated skills. Positive reinforcement, specific feedback and empathy were executed close to ‘extremely well’ by the prevention workers. This indicates that they focused on a positive group atmosphere.

We observed a high level of participant responsiveness, the third aspect of program integrity. The selection procedure may have contributed to this, as children’s motivation was checked before inclusion in the program (Kösters et al., 2012). Positively, children with higher initial levels of anxiety – and therefore most in need of the intervention – showed a higher degree of responsiveness.

Parental attendance was not well registered, but prevention workers indicated that their attendance was poor, a problem often reported in prevention research (Neil & Christensen, 2009). Children attended most of the sessions. This is most likely a positive consequence of the implementation during school time: attendance was not hampered by extracurricular activities and it did not demand time and effort of parents to bring the children to the sessions.

Children appraised FRIENDS for Life very positively: they found it useful and enjoyable, and almost all children indicated that they would recommend the program to other children, which is comparable with other FRIENDS for Life studies (e.g., Barrett, Shortt, Fox, & Wescombe, 2001; Gallegos-Guajardo et al., 2013; Stallard et al., 2005).

We hardly found any significant relations between program integrity and program outcomes, in contrast to reviews reporting that higher program integrity is related to more favorable outcomes (Dane & Schneider, 1998; Durlak & DuPre, 2008). Two aspects of quality of delivery were significantly associated with outcomes. Prevention workers’ self-disclosure was related to fewer self-reported anxiety symptoms at post-intervention, and specific feedback with fewer depression symptoms. However, these results have to be interpreted with caution, as some skills were performed at such a low rate that it was more likely that we measured the frequency rather than the quality.
of these skills. Although the checklist used to assess quality of delivery was designed specifically for FRIENDS for Life (Barrett, 1999), it may have been more appropriate to observe treatment sessions instead of prevention sessions, as the less frequently observed skills, paraphrasing, summarization, and reflection, are interviewing and counseling skills (Ivey & Ivey, 1994).

As the registration of homework was incomplete, we could not test the relation between homework on program outcomes. However, checking last session’s homework was an important part of every session. Children’s FRIENDS for Life workbook contained a homework reward card, and children received a sticker for every homework exercise they made. When they received a certain amount of stickers, the children could choose a gift (e.g., a small toy). Children were very eager to collect these. In addition, we observed that the prevention workers were very strict regarding homework. If children had not finished their homework, they had to show it the next week. Therefore, we may assume that most of the homework exercises were executed.

We found that children who appraised the program as more useful reported fewer anxiety symptoms at T2. However, as appraisal and outcomes were assessed at the same time, we cannot establish directionality between these measures.

Even though we found that prevention workers did not adhere completely to the protocol in the present study, this does not seem to affect program outcomes negatively: in the current trial, intervention group children with elevated anxiety and depression scores reported scores as low as the general population’s levels at 12 months post-intervention (Kösters et al., 2015). Presumably, the prevention workers chose to attune the program messages to the groups’ needs. Because of prevention workers’ experience in this particular setting, they may have delivered the program message without adhering completely. This seemed to have a positive effect on children’s appraisal, as children appraised the program more positively when the adherence to protocol was lower. This corresponds with what we noticed during the observations. For example, changing a reading exercise into a role play was very much welcomed by the children.

While higher adherence to protocol is associated with better program outcomes (Dane & Schneider, 1998; Durlak & DuPre, 2008), positive effects of adaptation have been reported previously (2008). The ideal balance between adherence and adaptation has yet to be established (Durlak & DuPre, 2008). A difficulty concerning CBT interventions is that it is still not clear what the effective mechanisms are (James, James, Cowdrey, Soler, & Choke, 2013), and therefore it remains indefinite which components should at least be preserved. Only for the CBT component ‘homework’, to which much attention had been given in the present setting, associations with better program outcomes have been found (Sandler et al., 2014). Up till now, no clear evidence has been found for the effectiveness of parent and booster sessions (Neil & Christensen, 2009; Sandler et al., 2014), of which prevention workers in the current study implemented less than the protocol prescribed. The positive findings on anxiety and depression reduction in the current trial (Kösters et al., 2015), indeed suggest that these sessions are not the most important components of anxiety and depression prevention programs.

The present study has several strengths. Firstly, we examined the implementation
of FRIENDS for Life in a naturalistic setting. This is important, as once a prevention program is disseminated on a larger scale, the implementation will be less controlled than in a research setting. Secondly, our study is, to our knowledge, the first to assess the relation between program integrity and program outcomes of FRIENDS for life as a preventive intervention. Contrary to previous studies, we measured four aspects of program integrity simultaneously instead of one or two (e.g., Essau et al., 2012; Barrett & Turner, 2001; Miller et al., 2011). Moreover, our study is the first that studied program integrity and children’s appraisal extensively for FRIENDS for Life as an indicated prevention program. Fourthly, independent observers assessed program integrity, minimizing social desirable reports (Dane & Schneider, 1998).

The present study has limitations as well. Although live observations provided a rich experience of the sessions, it was not possible to replay the sessions until every detail had been given full attention, as would have been possible if we had recorded the sessions. During the live observations, up to 11 children and two prevention workers had to be observed simultaneously. Therefore, details may have slipped from our attention. In addition, we cannot rule out that our presence during the observations influenced the behavior of the prevention workers or children.

CONCLUSION

In the present study, we investigated the implementation of FRIENDS for Life, a protocolled indicated preventive intervention, when executed in a naturalistic setting. We found that trainers applied positive skills, and that children’s responsiveness and attendance were good. Although the program was not completely executed according to the protocol, this seemed not to affect program outcomes negatively. Adaptation even seemed to have a positive influence, as children appraised the program more positively when trainers adhered less to the protocol. Future research should address the effects of adaptation on program outcomes.

IMPLICATIONS FOR PRACTICE

A highly protocolled intervention can be successfully transferred into daily school practice. A certain degree of adaptation by experienced mental health professionals does not necessarily have negative implications for program outcomes.
REFERENCE LIST


Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of


An evaluation of the FRIENDS programme: A cognitive behaviour therapy intervention to promote emotional resilience. Archives of Disease in Childhood, 90(10), 1016-1019.


