

Long-Term Outcome of a Cluster-Randomized Universal Prevention Trial Targeting Anxiety and Depression in School Children

Johan Ahlen*

Uppsala University

Fabian Lenhard

Ata Ghaderi

Karolinska Institutet

The present study concerns a 3-year follow-up of a universal prevention trial targeting anxiety and depressive symptoms in school children. In addition to evaluating the long-term effect of the prevention program, we also examined attrition and its effect on the outcome. High rates of attrition have commonly been observed in studies in the field. However, the role of attrition is not sufficiently understood regarding internal and external validity biases. The current study comprised 695 children (aged 8–11 at baseline) from 17 schools in Sweden. Schools were cluster-randomized to either the intervention or control condition. Children completed measures of anxiety and depressive symptoms and parents completed measures of their child's anxiety and general mental health. We found no evidence of long-term effects of the prevention program, except for a small effect regarding parent reports of child anxiety. However, that effect was not found to be of clinical significance. Regarding attrition, children with missing data at the 3-year follow-up displayed higher levels of psychiatric symptoms at baseline and increasing symptoms across time. Furthermore, children in the control condition with missing follow-up data were found to be significantly deteriorated across time compared to the corresponding children in the intervention condition regarding depressive symptoms and total difficulties. In

other words, attrition served as a moderator of the effect, which suggests that the overall result was biased toward a null-result. Our study highlights that large and nonrandom attrition severely limits the validity of the results. Further, given the common problem of retaining participants in long-term evaluations of school-based prevention trials, previous studies may suffer from the same limitations as the current study.

Keywords: universal prevention; anxiety; depression; long-term effect; attrition

ANXIETY DISORDERS AND DEPRESSION are common disorders in elementary school children, with a prevalence of 12.3% and 3.7%, respectively (Costello, Egger, Copeland, Erkanli, & Angold, 2011; Merikangas et al., 2010, respectively). Anxiety disorders and depression have a severe negative impact on several life areas and have been shown to lead to future psychiatric disorders (Copeland, Shanahan, Costello, & Angold, 2009). Furthermore, anxiety disorders and depression involve high costs for society (Snell et al., 2013). For example, Bodden, Dirksen, and Bögels (2008) found a twentyfold increased societal cost of families with clinically anxious children compared to controls, mostly due to parental productivity losses (i.e., absence of work). Given these findings, Bodden et al. (2008) suggested that the societal costs caused by childhood anxiety disorders may be similar to the costs of childhood conduct disorders

Address correspondence to Dr. Johan Ahlen, Department of Psychology, Uppsala University, Box 1225, 751 42 Uppsala, Sweden; e-mail: johan.ahlen@psyk.uu.se

and possibly even higher than the costs of autism spectrum disorders.

Despite the high prevalence and severe consequences of childhood anxiety disorders and depression, only a minority of children with these disorders are identified and receive adequate care (Essau, 2005). This highlights the need for further development and evaluation of effective preventive interventions. Prevention of mental disorders is commonly classified into universal prevention (no selection of individuals), selective prevention (individuals are selected due to being exposed to a risk factor), or indicated prevention (individuals are selected based on increased symptoms; Mrazek & Haggerty, 1994). Although selective and indicated prevention typically involve stronger effects compared to universal prevention (e.g., Sanchez et al., 2018), some researchers argue for the benefits of using universal prevention. Besides the benefit of reaching all children in a certain context, universal prevention is assumed to be more easily integrated into the school curriculum, associated with low dropout rates, and may be a way to avoid the risk of stigmatization when selecting certain children (Fazel, Hoagwood, Stephan, & Ford, 2014). However, long-term evaluations of school-based universal prevention of anxiety and depression are rare, and thus, we have limited knowledge about the effects that can unfold from these programs over the course of several years. Moreover, previous long-term evaluations have suffered from large attrition rates, and have most commonly not evaluated the effect of attrition on the outcome, which has limited the possibility to draw valid conclusions (e.g., Gillham et al., 2007; Johnstone, Rooney, Hassan, & Kane, 2014).

To our knowledge, only three cluster-randomized trials (Barrett, Farrell, Ollendick, & Dadds, 2006; Johnstone et al., 2014; Spence, Sheffield, & Donovan, 2005), and one randomized trial (Gillham et al., 2007) of school-based universal prevention of anxiety or depression have included longer follow-ups than 2 years. The study by Barrett et al. (2006) evaluated the FRIENDS for Life (FFL), a program originally developed to prevent anxiety in children. In the original study, significantly lower levels of anxiety and depressive symptoms were found in the intervention condition compared to the control condition at post and at the 1-year follow-up (Lock & Barrett, 2003). Further, at the 3-year follow-up, children in primary school who had received the FFL showed significantly lower levels of anxiety symptoms compared to the control group (Barrett et al., 2006). However, to the best of our knowledge, the analyses did not control for differences in the

baseline scores. This makes the result difficult to interpret, as there appeared to be lower levels of anxiety and depressive symptoms in the intervention group at baseline (see Lock & Barrett, 2003). The remaining three long-term follow-up studies (Gillham et al., 2007; Johnstone et al., 2014; Spence et al., 2005) primarily evaluated prevention of depressive symptoms. The study by Gillham et al. (2007) compared the Penn Resilience Program (PRP) to a control group. No overall effects were found regarding depressive symptoms neither at postassessment, nor at the 1-, 2-, or 3-year follow-up assessments ($d = 0.01$ at the 3-year follow-up). The study by Johnstone et al. (2014), comparing the Aussie Optimist Program (AOP) to a control group, found significantly reduced depressive symptoms at postintervention. However, the effect was not sustained at the 6-month follow-up or at subsequent follow-ups up to 4.5 years ($d = -0.13$ at the 4.5-year follow-up). Finally, the trial by Spence et al. (2005), evaluating the Problem Solving for Life program compared to a control group, found significantly reduced depressive symptoms at postintervention. However, the effect was not sustained at subsequent follow-ups over a 4-year period ($d = 0.05$ at the 4-year follow-up).

In summary, previous research on school-based universal prevention of anxiety and depression provides little evidence of its long-term effectiveness. A weakness of previous evaluations concerns the difficulties in retaining participants in long-term follow-ups, which have caused reduced study power, and consequently limited the possibility to detect any significant effects. In the trial by Spence et al. (2005) 61% of the sample was retained at the 4-year follow-up, in the trial by Barrett et al. (2006) 46% of the sample was retained at the 3-year follow-up, in the trial by Gillham et al. (2007) 43% of the sample was retained at the 3-year follow-up, and finally, in the trial by Johnstone et al. (2014) only 20% of the sample was retained at the 4.5-year follow-up. In addition to the increase of risk of type-II error that is associated with smaller samples, large attrition rates additionally limit the possibilities to make valid estimates of the intervention effects, because of threats to the external and internal validity. The external validity is compromised when the probability of attrition is related to sample characteristics such as gender, and the baseline level of the outcome variables (Foster & Fang, 2004). For example, if children with missing data at follow-up have higher levels of the outcome at baseline compared to completers, the external validity is compromised because the result is mainly valid for a subpopulation of the sample (i.e., those with lower baseline values). Even more

troublesome is the threat to the internal validity, which is likely to occur in cases where the probability of attrition is related to the outcome, and the strength of this relationship differs between the conditions (Foster & Fang, 2004). For example, it would be a threat to the internal validity if the association between parents' educational level and attrition was different between the conditions (e.g., parents with low educational level showed larger attrition compared to parents with high educational level in the control condition only, or vice versa). Specifically, it could both represent an inflated effect (i.e., if children with less educated parents to a larger degree benefited from the intervention), or conversely, the result could be biased towards the null-hypothesis (i.e., if children with better educated parents to a larger degree benefitted from the intervention).

In addition to high attrition rates in previous long-term evaluations, the effect of attrition has not been thoroughly addressed. Typically, analyses were conducted whether children with missing data at follow-up differ to completers at baseline (Barrett et al., 2006; Gillham et al., 2007; Johnstone et al., 2014; Spence et al., 2005). However, whether such differences were independent of the conditions has seldom been evaluated (i.e., only in Gillham et al., 2007). The lack of such analyses may increase the risk of misinterpretation of the result. Moreover, the effect of missing data on the outcome across time has only been evaluated in one study (i.e., in Spence et al., 2005). Such analyses are especially interesting as they can provide insights on different trajectories across time for children with missing data in the intervention condition compared to those with missing data the control condition.

Finally, to our knowledge, no previous study of school-based universal prevention of anxiety and depression has included parent ratings in the long-term follow-up, with one exception (Rooney et al., 2013), albeit with a large attrition rate (66%). Collecting information from multiple informants is important to comprehensively understand the effects. For example, in treatment studies of childhood anxiety disorders, effect sizes from self-reports have been smaller compared to parent's reports of child anxiety (Ishikawa, Okajima, Matsuoka, & Sakano, 2007).

THE ORIGINAL STUDY

Our original trial (Ahlen, Hursti, Tanner, Tokay, & Ghaderi, 2018) comprised 695 children, ages 8–11, cluster-randomized to either a teacher-administered school-based universal prevention program (FFL), or to a control condition. In the original study, we

found no evidence of an overall effectiveness of FFL, neither regarding anxiety nor depressive symptoms. However, in line with a recent meta-analysis (Sanchez et al., 2018), which found larger effect sizes in indicated prevention trials, we found an enhanced effect of FFL in children with elevated depressive symptoms at baseline ($d = 0.67$). Furthermore, we found decreased anxiety symptoms among children whose teachers attended a larger number of supervision sessions compared to children whose teachers attended fewer supervised sessions ($d = 0.22$) or compared to the control condition ($d = 0.21$).

THE PRESENT STUDY

The present study contributes to the current body of knowledge by investigating the long-term effects, potential moderators of outcome, and the influence of attrition on internal and external validity. The first aim of the study was to evaluate the long-term effects of a school-based universal prevention program (targeting anxiety and depression) 3 years after the completion of the program. The second aim was to examine the possible long-term maintenance of outcome in children with elevated depressive symptoms at baseline and children whose teachers attended the supervision to a high degree. Finally, the third aim was to evaluate differential patterns of attrition between conditions and the effect of attrition on the long-term outcome.

Method

The original study sample comprised 695 children (337 girls, 48%, and 358 boys, 52%) from 17 urban and suburban schools in Stockholm, Sweden, between the ages of 8–11 years ($M = 9.6$). The original study included assessments at baseline, postintervention, and at 1-year follow-up completed by children, parents, and teachers. Teachers in the intervention condition received a 1-day intense standardized training and subsequently administered the FFL 60 minutes per week, for 10 consecutive weeks during regular school hours. Teachers were also offered supervision with a clinical psychologist at three occasions during the implementation of the intervention (i.e., planning for future sessions and discussing potential obstacles and difficulties). Teachers in the control condition were instructed to run classes as usual. The evaluated program (FFL) is a cognitive and behavioral prevention program created to prevent the development of anxiety and depressive disorders in children. In short, children learn about recognizing and sharing feelings, about bodily clues of different emotions and different forms of relaxation, about helpful vs. unhelpful thoughts,

and about how they can overcome different problems by challenging situations using small steps, identifying a social support team, and using structured problem-solving. Additional details of participants and procedures are found in the original article by Ahlen et al. (2018). The trial was registered at [Clinicaltrials.gov](https://clinicaltrials.gov), trial identifying number NCT02134730. A full trial protocol can be offered upon request from the first author. The main funder of the trial was Uppsala University, Sweden. Funding covered education of teachers, assessments, and the planning and implementation of the trial by a doctoral student.

PARTICIPANTS

Figure 1 shows a flow chart of participants through each stage of the trial. Only 333 children (48%) still went to schools involved in the original study. To increase response rates, we contacted six additional secondary schools to which 158 children (23%) had been transferred. We also contacted 144 children (21%) by regular mail. The final sample at the 3-year follow-up consisted of 499 children (72% of the original sample). Most children ($n = 447$) completed the 3-year assessment in their classrooms and the remaining ($n = 52$) in their homes, in December 2016. A total of 30 children (4%) declined to participate, 14 children (2%) were absent from school at the assessment, 92 children (13%) did not respond to the invitation by mail, and 60 children (9%) were not reached with any information regarding the 3-year follow-up assessment. A total of 336 parents (48%) completed the 3-year follow-up via Internet in December 2016, 273 parents (40%) did not respond to the invitation, and 86 parents (12%) were not reached.

MEASURES

The Spence Children's Anxiety Scale (SCAS; Spence, 1998) is a measure of child anxiety symptoms. Excellent internal consistency of the total scores ($\alpha = .93$) was found in a Swedish sample together with evidence of convergent and divergent validity (Essau, Sasagawa, Anastassiou-Hadjicharalambous, Guzmán, & Ollendick, 2011). Internal consistency of total scores in our sample at the 3-year follow-up assessment was .92.

A parent-version of the SCAS is also available (SCAS-P, Spence, 1999). Good internal consistency of total scores ($\alpha = .89$) together with evidence of convergent and divergent validity was found in a Dutch sample (Nauta et al., 2004). Internal consistency of total scores in our sample at the 3-year follow-up assessment was .93.

The Children's Depression Inventory–Short Version (CDI-S; Kovacs, 2003) is a 10-item abbreviated form of the original Children's Depression Inventory. Good internal consistency of total scores ($\alpha = .80$) together with evidence of convergent validity was found in a Swedish sample (Ahlen & Ghaderi, 2017). Internal consistency of total scores in our sample at the 3-year follow-up assessment was .87.

The Strength and Difficulties Questionnaire (SDQ; Goodman, 1997) is a questionnaire developed to measure different aspects of children's mental health. The SDQ includes a measure of total difficulties (i.e., conduct problems, emotional problems, peer problems, and hyperactivity-inattention), and a measure of pro-social behaviors. Good internal consistency of total difficulties scores ($\alpha = .84$), acceptable internal consistency of pro-social scores ($\alpha = .67$) and evidence of discriminant validity were found in a Swedish sample (Malmberg, Rydell, & Smedje, 2003). Internal consistency of total difficulties scores in our sample at the 3-year follow-up assessment was .85, and the internal consistency of pro-social scores was .77.

PROCEDURE

The present study examined the results of a 3-year follow-up completed by children and parents. As agreed at the commencement, all schools in the control condition were offered to implement the FFL after the 1-year follow-up, including a full-day training for teachers. However, only four schools in the control condition agreed to proceed with the training of teachers, and unfortunately, none of the schools decided to implement the FFL after the 1-year follow-up. Reasons were several. Mainly, schools referred to new engagements in other projects, but some schools also expressed less enthusiasm because of the modest results found at the 1-year follow-up. Permission to conduct a 3-year follow-up was granted by the Regional Ethical Review board, after an amendment (Dnr: 2012/432/1) as this long-term follow-up was not originally planned in our project. We obtained an additional passive informed consent from parents, and an active informed consent from children. As the FFL was originally designed to prevent anxiety, the total score of the SCAS, and the SCAS-P served as primary outcomes, and the scores of CDI and SDQ served as secondary outcomes.

DATA ANALYSIS

All statistical analyses were performed in the R software program (R Core Team, 2015). We used linear mixed models (LMMs) to analyze the effects of the FFL. LMMs effectively models the dependent

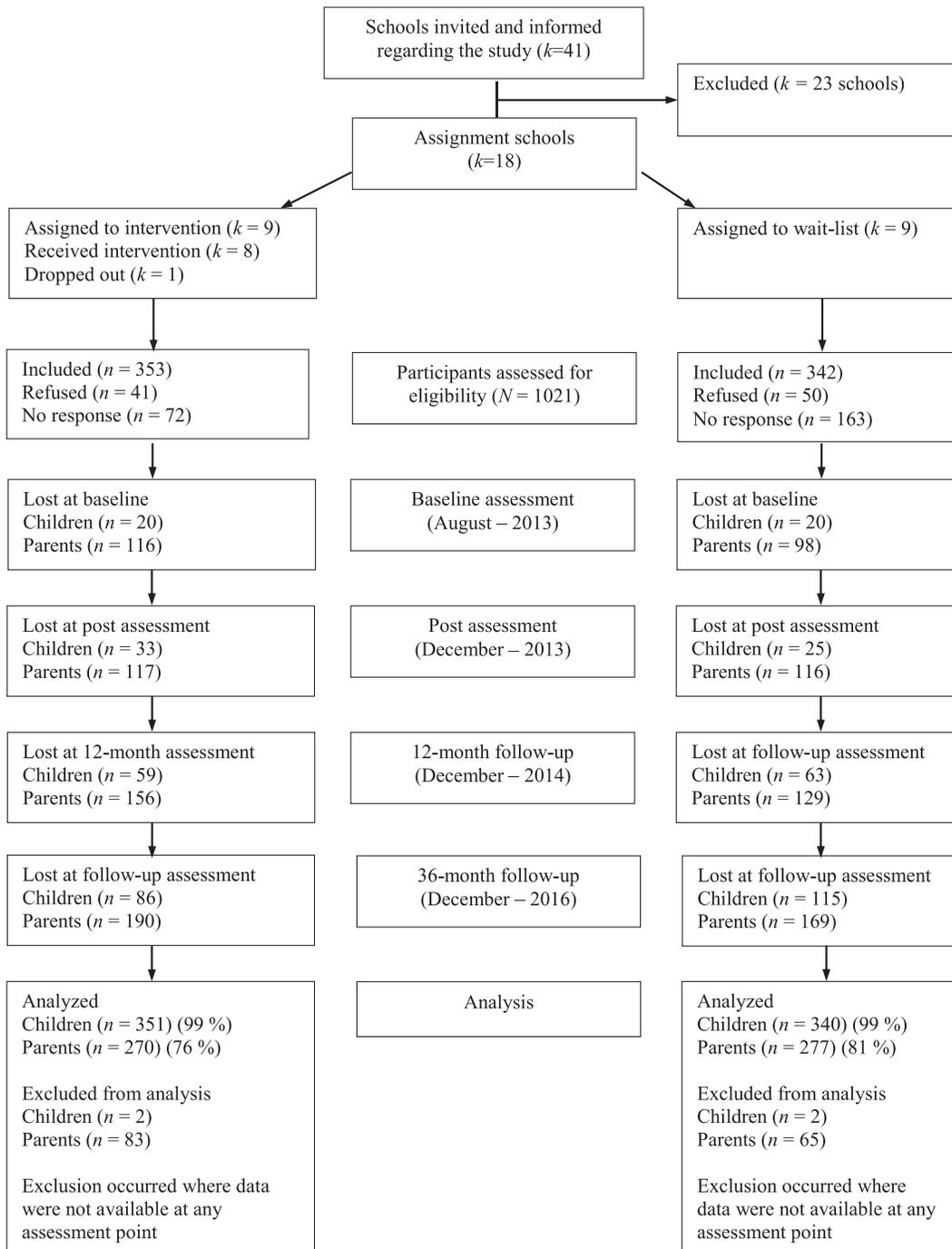


FIGURE 1 Flow of participants through each stage of the trial.

structure of the data by the inclusion of random effects, and involves the benefit of using all available data (i.e., LMMs include participants with partly missing observations; Hedeker & Gibbons, 1997). Thus, the LMM gives more weight to participants with more data points and do not delete participants with partly missing observations as in standard repeated measures ANOVA. To adequately model the study sample, we used a four-

level model with observations nested within subjects, subjects nested within classes, and classes nested within schools. The time variable was coded as number of months from baseline assessment. As often observed when using symptom questionnaires in a community sample, most outcomes were positively skewed. We therefore estimated bootstrapped confidence intervals of the fixed effects. For significant effects, we also examined the

Reliable Change Index (RCI; Jacobson & Truax, 1991) to interpret the clinical significance of the effects.

To analyze whether the short-term effect on depressive symptom in children with elevated baseline depressive symptoms was sustained at the 3-year follow up, we examined the baseline-symptoms*condition interaction effect in a LMM using changes in CDI-scores (from baseline to the 3-year follow-up) as the dependent variable. When analyzing the maintenance of the short-term effect on anxiety in relation to levels of teacher supervision, we divided the study sample into three groups: the control condition, the portion of the intervention condition whose teachers did not attend or only attended one session of supervision (low supervision group), and the portion of the intervention condition whose teachers attended two or all the three sessions of supervision (high supervision group). We then examined the supervision-group*time interaction effect in a LMM with self-rated anxiety as the dependent variable.

Regarding attrition, we followed the guidelines presented by Hedeker and Gibbons (1997). First, we coded two variables of missingness. One dummy-coded variable differentiating children that did not complete the 3-year assessment from those who did, and one variable differentiating children with missing parent-ratings from children with completed parent-ratings at the 3-year follow-up. Second, we examined missingness in relation to baseline symptoms in the full sample and in the two conditions separately in a series of Welch two sample t-tests (for continuous variables) and Fisher

exact tests (for categorical variables). Finally, we entered the variables of missingness in the LMMs examining the preventive effect, to evaluate any possible relationship between missingness and the outcome across time.

Results

LONG-TERM EFFECTS

Table 1 displays descriptive statistics for the outcome variables at all assessment points by condition. Table 2 presents detailed results from a series of LMMs evaluating the long-term effects. First, an LMM showed no significant main effect of condition (intervention vs. control) or time (months). Further, we found no significant condition*time interaction effect over the follow-up period on children's self-ratings of anxiety. Second, regarding parent reports of child anxiety, an LMM showed no main effect of condition or time. However, we found a significant condition*time interaction effect showing that children in the intervention condition displayed decreased levels of anxiety over time compared to the control condition (see Figure 2). In terms of the RCI, 9% were improved, 80% unchanged, and 11% deteriorated from baseline to the 3-year follow-up assessment in the intervention condition. Correspondingly, 12% were improved, 76% unchanged, and 12% deteriorated in the control condition. A Fisher's exact test revealed no significant difference between conditions on the RCI, $p = .727$. Third, regarding depressive symptoms, an LMM showed a significant main effect of time, meaning that overall, depressive symptoms increased over time. No main

Table 1

Means, Standard Deviations, and Number of Participants for Pre, Post, 1-year, and 3-year Follow-up Assessments Broken Down per Condition

Time-point	Intervention								Control							
	Pre		Post		1-year		3-year		Pre		Post		1-year		3-year	
	<i>M</i> (<i>SD</i>)	<i>n</i>														
SCAS	26.6 (15.7)	333	21.0 (15.1)	320	20.5 (13.5)	294	23.3 (15.0)	267	27.3 (14.4)	322	21.8 (15.8)	317	20.8 (13.5)	279	23.4 (14.0)	227
CDI-S	1.8 (2.5)	329	1.7 (2.5)	315	1.6 (2.5)	292	2.6 (3.3)	265	1.8 (2.5)	322	2.0 (3.1)	310	1.6 (2.5)	278	2.8 (3.5)	225
SCAS-P	15.5 (9.3)	237	15.1 (10.3)	236	15.4 (10.9)	197	15.0 (11.4)	163	14.6 (9.6)	244	13.0 (8.3)	226	13.9 (11.0)	213	14.5 (12.7)	173
SDQ-difficulties	7.0 (5.4)	232	7.5 (5.7)	235	7.4 (6.0)	193	7.1 (5.5)	161	6.1 (5.2)	241	6.5 (5.3)	226	6.3 (5.4)	213	6.3 (5.6)	171
SDQ-strengths	8.4 (1.8)	232	8.2 (2.0)	235	8.2 (1.8)	193	7.9 (1.9)	161	8.5 (1.5)	241	8.4 (1.6)	226	8.4 (1.6)	213	8.4 (1.8)	171

Note. SCAS = Spence Children's Anxiety Scale (child-ratings), CDI-S = Children's Depression Inventory, SCAS-P = Spence Children's Anxiety Scale Parent version, SDQ-difficulties = Strength and Difficulties Questionnaire Total difficulties subscale, SDQ-strengths = Strength and Difficulties Questionnaire – Pro-social behavior subscale.

Table 2
Main-, and Interaction-Effects in the Linear Mixed Models Evaluating the Friends for Life Intervention for All Outcomes

Outcome	Fixed effect	Coefficient (B)	95% CI	ES (d)	95% CI
SCAS	Group	-0.70	[-4.36, 2.99]	0.04	[-0.14, 0.21]
	Time (months)	-0.04	[-0.08, 0.01]	0.17	[-0.01, 0.34]
	Group*Time	0.004	[-0.05, 0.06]	-0.01	[-0.19, 0.16]
CDI-S	Group	-0.11	[-0.66, 0.42]	0.04	[-0.14, 0.22]
	Time (months)	0.03***	[0.02, 0.04]	-0.50	[-0.68, -0.32]
	Group*Time	-0.003	[-0.02, 0.01]	0.04	[-0.13, 0.22]
SCAS-P	Group	1.45	[-0.48, 3.62]	-0.16	[-0.37, 0.05]
	Time (months)	0.03	[-0.01, 0.05]	-0.19	[-0.40, 0.03]
	Group*Time	-0.04*	[-0.08, -0.01]	0.23	[0.02, 0.45]
SDQ-difficulties	Group	1.03	[-0.44, 2.29]	-0.18	[-0.40, 0.03]
	Time (months)	0.01	[-0.003, 0.02]	-0.15	[-0.36, 0.07]
	Group*Time	-0.01	[-0.03, 0.01]	0.15	[-0.07, 0.36]
SDQ-strengths	Group	-0.20	[-0.48, 0.12]	-0.15	[-0.37, 0.06]
	Time (months)	-0.004	[-0.01, 0.001]	-0.18	[-0.40, 0.03]
	Group*Time	-0.006	[-0.01, 0.001]	-0.16	[-0.37, 0.05]

Note. ES=Effect-sizes, are converted into Cohen's *d* from the *t*-value of the parameter estimate and are reported as positive when they appeared in the predicted direction (e.g. lower anxiety symptoms in the intervention condition, and lower anxiety symptoms over time). SCAS = Spence Children's Anxiety Scale (child-ratings), CDI-S = Children's Depression Inventory, SCAS-P = Spence Children's Anxiety Scale Parent version, SDQ-difficulties = Strength and Difficulties Questionnaire Total difficulties subscale, SDQ-Strengths = Strength and Difficulties Questionnaire – Pro-social behavior subscale.

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

effect of condition, or condition*time interaction effect was found on depressive symptoms. Finally, no significant main effects of time or condition, or condition*time interaction effect was found on parent reports of child difficulties and pro-social behaviors.

MAINTENANCE OF SHORT-TERM EFFECTS

Regarding children with elevated depressive symptoms at baseline, an LMM showed no significant baseline-symptom*condition interaction effect on depressive symptoms over the 3-year follow-up period, $B=0.04$, 95% CI [-0.25, 0.34], $d=-0.03$, 95% CI [-0.24, 0.18]. Further, an LMM showed no significant long-term reduction in anxiety symptoms in the high supervision group compared to the control condition, $B=0.01$, 95% CI [-0.06, 0.08], $d=-0.02$, [-0.27, 0.22], or the low supervision group, $B=0.01$, 95% CI [-0.06, 0.08], $d=-0.04$, [-0.33, 0.24]. Consequently, the short-term effects found post intervention were not evident over the 3-year follow-up period.

CHILD ATTRITION

Table 3 presents baseline descriptive statistics by missingness. Table 4 presents the corresponding descriptive statistics separately for the intervention

and control condition. In the result and discussion section, we refer to child attrition when describing children with missing self-ratings, and parent attrition when describing children with missing parent-ratings at the 3-year follow-up.

A Welch two sample *t*-test showed that overall, children with missing data at the 3-year follow-up had significantly higher parent reports of child anxiety and total difficulties at baseline assessment, $t(208.04)=2.11$, $p=0.036$, $d=0.23$; $t(214.07)=4.05$, $p<.001$, $d=0.44$, respectively (see Table 3). We found no such differences at baseline assessment regarding symptoms of self-rated anxiety and depressive symptoms, parent reports of child pro-social behaviors, age, gender, household income, or parent's educational level.

Regarding different patterns of attrition between conditions (see Table 4), a Fisher exact test showed that a higher proportion of children were missing at the 3-year follow-up in the control condition (33%) compared to the intervention condition (24%), $p=.009$. Further, in the control condition (but not in the intervention) children missing at the 3-year follow-up displayed higher parent reports of child anxiety, $t(117.69)=2.39$, $p=.018$, $d=0.36$. Finally, in the control condition, children missing at the 3-year follow-up were older than children who

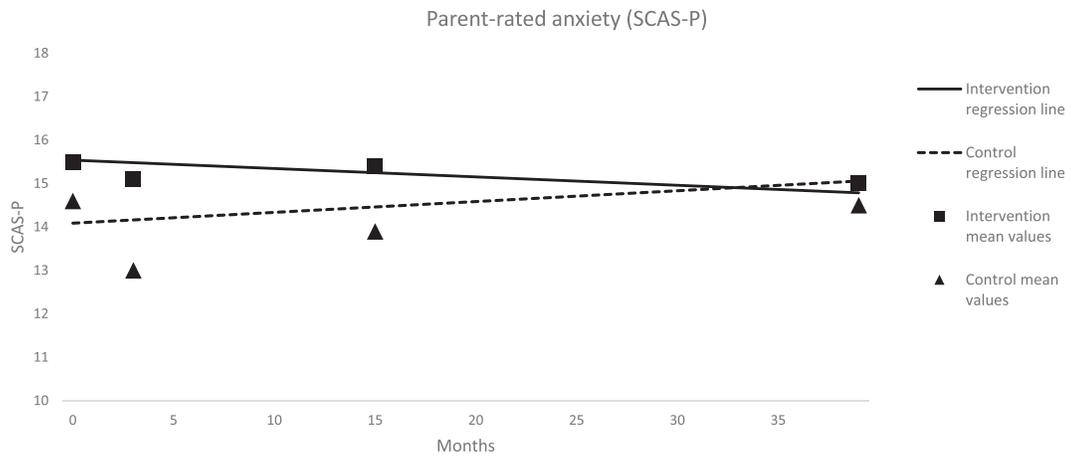


FIGURE 2 Estimated regression lines for the intervention and control condition separately, and mean values (based on raw data) at all assessment points for the intervention and control condition separately. SCAS-P = Spence Children's Anxiety Scale Parent version.

completed the 3-year follow-up, $t(187.04)=2.54$, $p=.012$, $d=0.31$.

PARENT ATTRITION

Children with missing parent ratings at the 3-year follow-up had significantly higher self-rated anxiety, $t(651.48)=2.18$, $p=.029$, $d=0.17$, depressive symptoms, $t(628.92)=3.09$, $p=.002$, $d=0.25$, and lower household income, $t(337.20)=3.22$, $p=.001$, $d=0.31$ (see Table 3). We found no significant difference in attrition rates between conditions (see Table 4: intervention=54%, control=49%). The difference in lower household income was only observed in the intervention condition when

running the analysis separately for each condition, $t(164.13)=3.16$, $p=.002$, $d=0.44$.

CHILD ATTRITION AND ITS EFFECTS ON THE OUTCOME

When adding the variable of child attrition into the LMMs examining group and time effects on self-rated anxiety, parent ratings of child anxiety and total difficulties, we found no main or interaction effects of missingness on the outcome. However, when adding the variable of child attrition into the LMM examining depressive symptoms, we found a main effect of child attrition, $B = 0.72$, 95% CI [0.12, 1.31], $d = 0.19$, meaning that overall,

Table 3

Baseline Descriptive Statistics by Missingness Regarding Child Attrition (Missing Child Ratings at 3-Year Follow-up), and Parent Attrition (Missing Child Ratings at 3-Year Follow-up)

Baseline variables	Child attrition		Parent attrition	
	Missing	Completers	Missing	Completers
Number of children (%)	196 (28%)	499 (72%)	359 (52%)	336 (48%)
Mean age	9.6 years	9.6 years	9.6 years	9.6 years
Percentage girls	48%	48%	50%	46%
Higher education ^a	56%	57%	53%	58%
Mean household income ^b	6000 USD	6600 USD	5500 USD	6600 USD
Mean SCAS (<i>SD</i>)	26.6 (15.1)	27.1 (15.1)	28.1 (16.0)	25.6 (13.9)
Mean CDI-S (<i>SD</i>)	2.1 (3.0)	1.7 (2.3)	2.1 (2.8)	1.5 (2.1)
Mean SCAS-P (<i>SD</i>)	16.6 (10.8)	14.4 (8.8)	15.1 (10.0)	15.0 (9.1)
Mean SDQ-difficulties (<i>SD</i>)	8.2 (5.9)	5.9 (4.9)	7.2 (5.8)	6.2 (5.1)
Mean SDQ-strengths (<i>SD</i>)	8.4 (1.6)	8.5 (1.7)	8.4 (1.7)	8.5 (1.7)

Note. ^a Parents with a post-secondary educational. ^b Mean household income per month. SCAS = Spence Children's Anxiety Scale (child-ratings), CDI-S = Children's Depression Inventory, SCAS-P = Spence Children's Anxiety Scale Parent version, SDQ-difficulties = Strength and Difficulties Questionnaire Total difficulties subscale, SDQ-Strengths = Strength and Difficulties Questionnaire – Pro-social behavior subscale.

Table 4

Baseline Descriptive Statistics by Condition and Missingness Regarding Child Attrition (Missing Child Ratings at 3-Year Follow-up), and Parent Attrition (Missing Child Ratings at 3-Year Follow-up)

Baseline variables	Child attrition				Parent attrition			
	Intervention group		Control group		Intervention group		Control group	
	Missing	Completers	Missing	Completers	Missing	Completers	Missing	Completers
Number of children (%)	84 (24%)	269 (76%)	112 (33%)	230 (67%)	190 (54%)	163 (46%)	169 (49%)	173 (51%)
Child's age	9.7	9.7	9.5	9.4	9.7	9.8	9.4	9.4
Percentage girls	46%	46%	50%	51%	47%	45%	54%	47%
Higher education ^a	50%	58%	60%	56%	52%	58%	54%	59%
Mean household income ^b	7200 USD	6600 USD	6000 USD	6000USD	6000 USD	7200 USD	5500 USD	6000 USD
Mean SCAS (<i>SD</i>)	25.7 (14.6)	26.9 (16.0)	27.2 (15.5)	27.3 (13.9)	28.0 (16.6)	24.9 (14.4)	28.3 (15.3)	26.2 (13.5)
Mean CDI-S (<i>SD</i>)	2.2 (3.2)	1.7 (2.2)	2.0 (2.9)	1.7 (2.3)	2.1 (2.8)	1.4 (2.1)	2.1 (2.8)	1.5 (2.2)
Mean SCAS-P (<i>SD</i>)	16.1 (10.4)	15.2 (9.0)	17.0 (11.2)	13.6 (8.5)	14.9 (9.7)	15.8 (9.1)	15.3 (10.2)	14.3 (9.2)
Mean SDQ-difficulties (<i>SD</i>)	8.9 (5.9)	6.4 (5.1)	7.7 (5.9)	5.4 (4.7)	7.3 (5.6)	6.9 (5.3)	7.1 (5.9)	5.6 (4.7)
Mean SDQ-strengths (<i>SD</i>)	8.4 (1.9)	8.4 (1.8)	8.4 (1.4)	8.6 (1.6)	8.4 (1.7)	8.3 (1.9)	8.4 (1.6)	8.6 (1.5)

Note. ^a Parents with a post-secondary educational. ^b Mean household income per month. SCAS = Spence Children's Anxiety Scale (child-ratings), CDI-S = Children's Depression Inventory, SCAS-P = Spence Children's Anxiety Scale Parent version, SDQ-difficulties = Strength and Difficulties Questionnaire Total difficulties subscale, SDQ-strengths = Strength and Difficulties Questionnaire – Pro-social behavior subscale.

children with missing data had higher depressive symptoms compared to children with complete data at the 3-year follow-up. When adding child attrition into the LMM of total difficulties, we found a main effect of child attrition, $B = 2.18$, 95% CI [0.78, 3.38], $d = 0.25$, meaning that over all, children missing at last follow-up had higher total difficulties as rated by parents. We also found a child attrition*time effect, $B = 0.03$, 95% CI [0.01,

0.07], $d = 0.15$, meaning that children missing at 3-year follow-up showed a relative increase in total difficulties over time. Lastly, we also found a child attrition*condition*time effect, $B = -0.06$, 95% CI [-0.01, -0.11], $d = 0.17$, meaning that missing children in the intervention condition had a relative decrease in difficulties over time compared to the corresponding children in the control condition (see Figure 3).

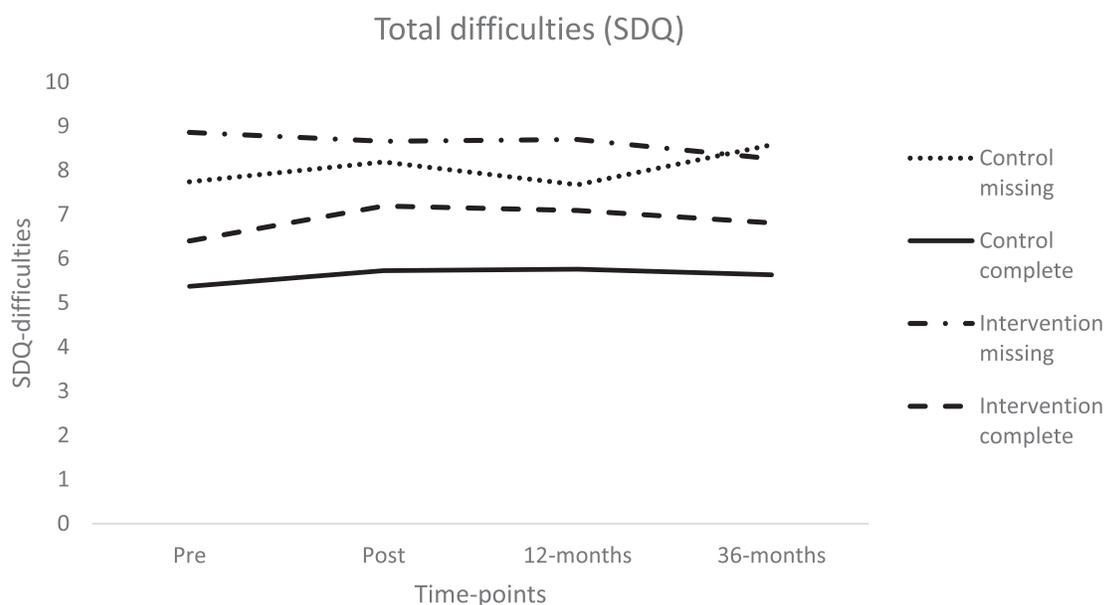


FIGURE 3 The course of total difficulties symptoms over time by condition and child attrition.

PARENT ATTRITION, EFFECTS ON THE OUTCOME

When adding the variable of parent attrition into the LMMs of self-rated anxiety, parent ratings of child anxiety and pro-social behavior, we found no main or interaction effects of missingness on the outcome. However, when adding the variable of parent attrition into the LMM of total difficulties, we found a main effect of parent attrition, $B = 1.43$, 95% CI [0.05, 2.77], $d = 0.16$, meaning that overall, children with missing parent ratings had higher total difficulties compared to children with complete parent ratings at the 3-year follow-up. Further, when adding parent attrition into the LMM of depressive symptoms, we found a parent attrition*time effect, $B = 0.03$, 95% CI [0.01, 0.05], $d = 0.23$, meaning that children with missing parent ratings had a relative increase in depressive symptoms over time compared to children with complete parent ratings at the 3-year follow-up. Finally, we found a parent attrition*condition*time effect, $B = -0.03$, 95% CI [-0.01, -0.05], $d = 0.16$, meaning that the increase in depressive symptoms observed in children with missing parent ratings were significantly smaller in the intervention condition compared to the control condition (see Figure 4).

Discussion

The objective of the present study was to investigate the long-term outcomes of a teacher-administered universal cognitive-behavioral prevention program, and the possible moderating effects of depressive symptom severity, and level of supervision. The

study also aimed at examining attrition and its effects on internal and external validity of the results.

We found no differences between the two conditions regarding the course of symptoms over time, except for parent reports of child anxiety. However, in the subsequent analyses of clinical significance, we found no evidence of a meaningful prevention effect. The results of the long-term effectiveness evaluation are therefore in line with previous studies, which all have failed to find support for long-term effects of school-based universal prevention of anxiety and depression (Barrett et al., 2006; Gillham et al., 2007; Johnstone et al., 2014; Spence et al., 2005). One possible hypothesis (that may explain these results) is that the present study, together with previous long-term evaluations, suffered from type-II error, meaning that due to overall high attrition, sample sizes were too small to detect differences between conditions. However, the effect sizes in the present study (and previous studies) seem to approach zero, and there is no obvious tendency for any positive (yet nonsignificant) prevention effect.

Another hypothesis is that school-based universal prevention as implemented to date does not include any measurable long-term effects. A possible problem inherent in existing school-based programs like the FFL may be that too much responsibility is put on the child to integrate the newly learned skills and strategies into their behavioral repertoire. In previous studies, teachers have typically received only 1 or 2 days of training

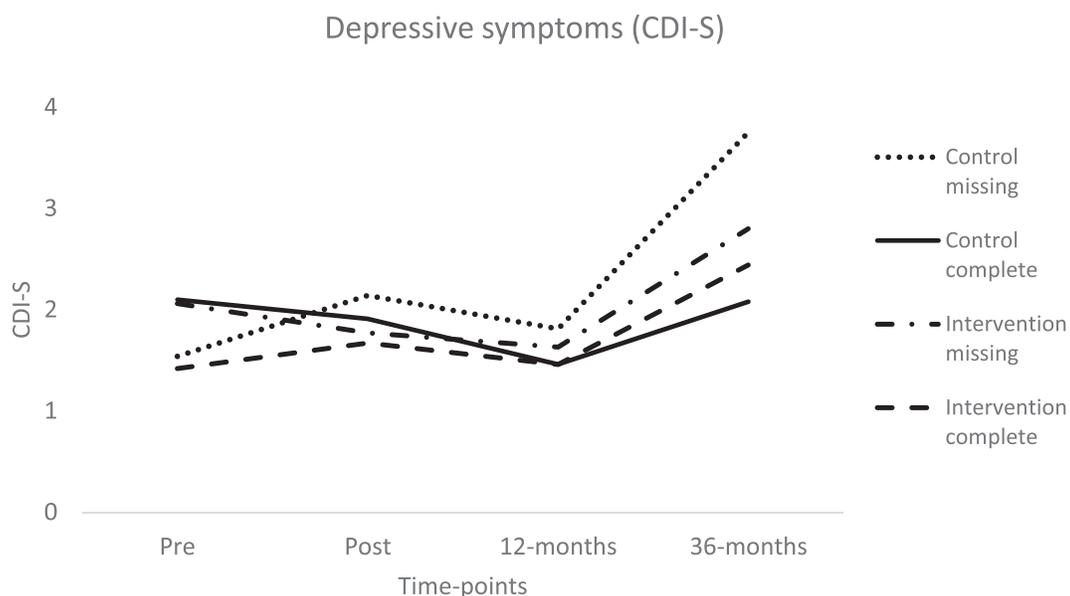


FIGURE 4 The course of depressive symptoms over time by condition and parent attrition.

to run school-based programs to prevent anxiety and depression (e.g., Barrett et al., 2006; Johnstone et al., 2014; Spence et al., 2005). Moreover, although many programs encourage parental involvement, previous school-based universal trials have typically described major problems in engaging parents in the program (Ahlen, Breitholtz, Barrett, & Gallegos, 2012; Lowry-Webster, Barrett, & Lock, 2003). To exemplify, a common strategy in these programs involves teaching the child how to modify unhelpful thoughts into helpful thoughts. An alternative approach would be to educate teachers to better recognize and reward helpful thoughts expressed by the child. Consequently, an approach where teachers are substantially trained in such core practice elements (for a discussion, see McLeod et al., 2017) instead of receiving the minimal training to run a manualized program may enhance the preventive effect. For example, in a recent meta-analysis of school-based mental health services, programs where strategies were integrated into the existing daily school activities showed larger effects compared to programs (such as the FFL) delivered in specific lessons added to the curriculum (Sanchez et al., 2018).

Further, we found no support of a long-term effect of the FFL for children whose teachers attended a large number of supervision sessions. This effect was only evident during the implementation of the sessions. These results might partly be explained by the general model within the field of school-based prevention to provide intensive training of teachers prior to the implementation, but often lack an extensive, continued support to secure high fidelity in the long run (Atkins, Cappella, Shernoff, Mehta, & Gustafson, 2017). In line with suggestions from Atkins et al. (2017), we recommend that future implementations of universal prevention programs include building support systems for teachers within the school in order to enhance the long-term sustainment of specific teacher skills that are crucial for adequate implementation of these programs. Specifically, previous research has highlighted the important role of principals and immediate colleagues in promoting adherence and support when implementing new strategies and programs in schools (Beets et al., 2008). For example, schools may benefit from creating so-called learning communities (for a discussion, see Vescio, Ross, & Adams, 2008) where teachers can reflect and discuss the implementation of the program with colleagues who share similar day-to-day experiences (Atkins et al., 2017).

When exploring the association between baseline characteristics and attrition in the intervention and

control condition separately, the analyses revealed different patterns of attrition between conditions. This was evident for both child and parent attrition, and both regarding baseline symptom levels and other demographic variables, implying that the equivalence of the groups was reduced and hence also the internal validity of the results. These results stand in contrast to the study of Gillham et al. (2007) where no different patterns between conditions were found despite higher attrition rates. One possible explanation of these differences might be due to different allocation strategies. More specifically, the present study used random assignment by clusters (i.e., schools), whereas the study by Gillham used random assignment at the individual level. Studies including cluster-randomization are more susceptible to clustered attrition, but have despite this (and its lower study power) often been the choice in school-based universal prevention trials (Ahlen, Lenhard, & Ghaderi, 2015). Probably, this choice has been due to practical reasons but also due to methodological reasons like reducing contamination effects (i.e., children in the control condition learn strategies by peers in the intervention condition). The present study points out an additional problem with cluster-randomization, as it might involve a higher vulnerability to condition-specific attrition.

Finally, our study suggests that attrition may have caused a wash out of an existing, true effect. The current study provides a few clues that give some preliminary support for this suggestion. For example, we found that children with missing data at the long-term follow-up had higher depressive symptoms and more total difficulties over the course of the study compared to children with complete data at the 3-year follow-up assessment. Consequently, this meant that high-burdened children were assigned less weight in the statistical analyses, and thus, children less burdened (with less room for improvement) were assigned more weight. Previous research has found that the prevention effect typically is amplified in subgroups of children with elevated symptoms (e.g., Ahlen et al., 2018; Spence et al., 2005) and that interventions that solely target children with elevated symptoms (i.e., indicated prevention) show larger effects compared to prevention delivered to a normal population (Sanchez et al., 2018). Accordingly, we would expect a reduced effect a less burdened subsample is assigned more weight. Moreover, and especially interesting in the present study, children with missing data at the 3-year follow-up were found to follow different trajectories over time dependent on condition. According to the parent ratings, we found a significant prevention effect on total

difficulties in children with missing follow-up data (self-reports). Correspondingly, according to the self-ratings, we found a significant prevention effect on depressive symptoms in children with missing parent ratings at follow-up. In other words, attrition served as a moderator of the effect. It is not impossible that washed out effects (due to high and nonrandom attrition) have been the case in previous long-term evaluations of school-based universal prevention. It is, for example, quite possible that children with emotional and behavioral difficulties to a larger degree drop out due to a higher degree of school mobility, as indicated by some studies (e.g., Rumberger, 2003).

FUTURE DIRECTIONS

The current study suggests that the long-term effects of universal schools-based prevention of anxiety and depression at present date are not sufficiently understood. The results indicate that the high attrition rates that usually are found in the field pose serious limitations to the interpretations of the results. We believe that there is a need to further evaluate the long-term effectiveness of comparable prevention programs, although future studies face many difficult challenges. First, as previous meta-analyses suggest, effect sizes of universal prevention of anxiety and depression are very small (Ahlen et al., 2015; Sanchez et al., 2018). To attain power to detect such small effects, future trials need to recruit larger sample sizes than those of the previous trials. For example, to find an effect size of 0.20 (Cohen's d), which is larger than the average between-group differences in universal prevention (Ahlen et al., 2015), a cluster-randomized trial with two arms would need to include $N > 1200$, i.e. more than 40 schools with 30 children from each school. Further, future studies need to address the problem of nonrandom attrition: for example, including strategies to reach children who switch schools during the follow-up period. For example, in the enrollment phase researchers should carefully collect several contact details to increase chances of tracking participants over time. Moreover, future studies also need to reduce the condition-specific attrition, which we suggest may be of particular concern in cluster-randomized trials. Researchers need to balance the decision of what randomization procedure to implement by carefully considering the pros and cons of each procedure. In the case of individual randomization, researchers need to estimate the risk of contamination and implementation problems such as increased resource demands for the school due to splitting classes during the intervention. Specifically associated with cluster-randomization, researchers

need to estimate the risk of clustered attrition (i.e., data missing not at random) and increased costs and demands for the research group due to larger sample sizes.

Beyond challenges of attrition, we believe future studies need to further consider program content and delivery of programs to strengthen the possibility of long-term effects. For example, schools-based universal prevention programs typically include numerous strategies aiming to target different risk factors involved in the development of anxiety and depression (Briesch, Sanetti, & Briesch, 2010; Brunwasser, Gillham, & Kim, 2009). However, at present, it is not evident which specific strategies are more effective than others in universal prevention of anxiety and depression (Ahlen et al., 2015; Sanchez et al., 2018). In treatment research, recent studies indicate that some parts typically included in treatment programs targeting anxiety and depression may not yield any additional effect (e.g., Kendall et al., 2016; Richards et al., 2016). For example, in the study by Kendall and colleagues (2016), the child's ability to cope with anxiety-provoking situations was found to be a mediator of the effect. In contrast, reductions in anxious self-talk were not found to mediate the effect of the treatment. A large number of the exercises in the FFL involve work to identify and modify negative self-talk. Perhaps a streamlined program that more clearly focuses on helping the child to cope with anxiety-provoking situations could be one way to achieve long-term effects.

Finally, future studies need to reflect on several implementation questions such as the amount of training and supervision to ensure high fidelity of the intervention, when delivered by the administrator (e.g., teachers). Recently, Owens and colleagues (2014) highlighted the importance of evaluating several research questions related to training and supervision in future prevention research. For example, what dosage of training and supervision is needed? What training and supervision strategies are most successful? And, do the administrators' attitudes about the program predict successful implementation? Beyond these questions, it is also important to discuss whether universal school-based prevention programs should be administered by mental health professionals or teachers. Although previous meta-analyses suggest that prevention administered by mental health professionals involves larger effects compared to prevention administered by teachers (Stice et al., 2009; Teubert & Pinquart, 2011), there is no clear answer to this question. For example, Stice and colleagues (2009) reported that indicated prevention trials typically involved mental health

professionals, while school-based prevention trials seldom do (Sanchez et al., 2018). Consequently, the differences in effects between studies using mental health professionals or teachers may be partially explained by the association with child's risk status. Further, as mentioned above, recent results also highlight possible superiority of interventions integrated to the daily school activities, which speaks in favor of using teachers (Sanchez et al., 2018).

LIMITATIONS

A major limitation of the current study is the relatively low power, which limits the possibility to draw any firm conclusions of the long-term effects. Although the current study retained a larger proportion of participants compared to previous studies, it suffered from nonrandom and condition-specific attrition, which severely affected interpretation of the results.

CONCLUSIONS

To conclude, the current study failed to find a clinically significant long-term effect of a universal prevention program targeting anxiety and depression. However, attrition was found to suppress the effect on self-rated depressive symptoms and total difficulties as rated by the parents. Previous long-term evaluations of school-based prevention of anxiety and depression may suffer from similar attrition effects, given that these studies have involved even larger attrition rates than the current study. Future long-term evaluations need to plan how to reach children and youths who switch schools during the follow-up period, as well as addressing the risk of condition-specific attrition in cluster-randomized trials.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

References

- Ahlen, J., Breitholtz, E., Barrett, P. M., & Gallegos, J. (2012). School-based prevention of anxiety and depression: A pilot study in Sweden. *Advances in School Mental Health Promotion, 5*, 246–257. <https://doi.org/10.1080/1754730X.2012.730352>
- Ahlen, J., & Ghaderi, A. (2017). Evaluation of the children's depression inventory—short version (CDI-S). *Psychological Assessment, 29*, 1157–1166. <https://doi.org/10.1037/pas0000419>
- Ahlen, J., Lenhard, F., & Ghaderi, A. (2015). Universal prevention for anxiety and depressive symptoms in children: A meta-analysis of randomized and cluster-randomized trials. *The Journal of Primary Prevention, 36*, 387–403. <https://doi.org/10.1007/s10935-015-0405-4>
- Ahlen, J., Hursti, T., Tanner, L., Tokay, Z., & Ghaderi, A. (2018). Prevention of anxiety and depression in Swedish school children: A cluster-randomized effectiveness study. *Prevention Science, 19*, 147–158. <https://doi.org/10.1007/s11121-017-0821-1>
- Atkins, M. S., Cappella, E., Shernoff, E. S., Mehta, T. G., & Gustafson, E. L. (2017). Schooling and children's mental health: Realigning resources to reduce disparities and advance public health. *Annual Review of Clinical Psychology, 13*, 123–147. <https://doi.org/10.1146/annurev-clinpsy-032816-045234>
- Barrett, P. M., Farrell, L. J., Ollendick, T. H., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: An evaluation of the friends program. *Journal of Clinical Child & Adolescent Psychology, 35*, 403–411. https://doi.org/10.1207/s15374424jccp3503_5
- Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K. K., & Allred, C. (2008). School climate and teachers' beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science, 9*, 264–275. <https://doi.org/10.1007/s11121-008-0100-2>
- Bodden, D. H., Dirksen, C. D., & Bögels, S. M. (2008). Societal burden of clinically anxious youth referred for treatment: A cost-of-illness study. *Journal of Abnormal Child Psychology, 36*, 487–497. <https://doi.org/10.1007/s10802-007-9194-4>
- Briesch, A. M., Sanetti, L. M. H., & Briesch, J. M. (2010). Reducing the prevalence of anxiety in children and adolescents: An evaluation of the evidence base for the FRIENDS for Life program. *School Mental Health, 2*, 155–165. <https://doi.org/10.1007/s12310-010-9042-5>
- Brunwasser, S. M., Gillham, J. E., & Kim, E. S. (2009). A metaanalytic review of the Penn Resiliency Program's effect on depressive symptoms. *Journal of Consulting and Clinical Psychology, 77*, 1042–1054. <https://doi.org/10.1037/a0017671>
- Copeland, W. E., Shanahan, L., Costello, E. J., & Angold, A. (2009). Childhood and adolescent psychiatric disorders as predictors of young adult disorders. *Archives of General Psychiatry, 66*, 764–772. <https://doi.org/10.1001/archgenpsychiatry.2009.85>
- Costello, E. J., Egger, H. L., Copeland, W., Erkanli, A., & Angold, A. (2011). The developmental epidemiology of anxiety disorders: Phenomenology, prevalence, and comorbidity. In W. K. Silverman, & A. P. Field (Eds.), *Anxiety disorders in children and adolescents* (pp. 56–75). (2nd ed.). New York, NY: Cambridge University Press.
- Essau, C. A. (2005). Frequency and patterns of mental health services utilization among adolescents with anxiety and depressive disorders. *Depression and Anxiety, 22*, 130–137. <https://doi.org/10.1002/da.20115>
- Essau, C. A., Sasagawa, S., Anastassiou-Hadjicharalambous, X., Guzmán, B. O., & Ollendick, T. H. (2011). Psychometric properties of the Spence Child Anxiety Scale with adolescents from five European countries. *Journal of Anxiety Disorders, 25*, 19–27. <https://doi.org/10.1016/j.janxdis.2010.07.001>
- Fazel, M., Hoagwood, K., Stephan, S., & Ford, T. (2014). Mental health interventions in schools in high-income countries. *The Lancet Psychiatry, 1*, 377–387. [https://doi.org/10.1016/S2215-0366\(14\)70312-8](https://doi.org/10.1016/S2215-0366(14)70312-8)
- Foster, E. M., & Fang, G. Y. (2004). Alternative methods for handling attrition: An illustration using data from the Fast Track evaluation. *Evaluation Review, 28*, 434–464. <https://doi.org/10.1177/0193841X04264662>
- Gillham, J. E., Reivich, K. J., Freres, D. R., Chaplin, T. M., Shatté, A. J., Samuels, B., & Seligman, M. E. (2007). School-

- based prevention of depressive symptoms: A randomized controlled study of the effectiveness and specificity of the Penn Resiliency Program. *Journal of Consulting and Clinical Psychology*, 75, 9–19. <https://doi.org/10.1037/0022-006X.75.1.9>
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586. <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>
- Hedeker, D., & Gibbons, R. D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, 2, 64–78. <https://doi.org/10.1037/1082-989X.2.1.64>
- Ishikawa, S. I., Okajima, I., Matsuoka, H., & Sakano, Y. (2007). Cognitive behavioural therapy for anxiety disorders in children and adolescents: A meta-analysis. *Child and Adolescent Mental Health*, 12, 164–172. <https://doi.org/10.1111/j.1475-3588.2006.00433.x>
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Johnstone, J., Rooney, R. M., Hassan, S., & Kane, R. T. (2014). Prevention of depression and anxiety symptoms in adolescents: 42 and 54 months follow-up of the Aussie Optimism Program Positive Thinking Skills. *Frontiers in Psychology*, 5, 1–10. <https://doi.org/10.3389/fpsyg.2014.00364>
- Kendall, P. C., Cummings, C. M., Villabø, M. A., Narayanan, M. K., Treadwell, K., Birmaher, B., & Gosch, E. (2016). Mediators of change in the Child/Adolescent Anxiety Multimodal Treatment Study. *Journal of Consulting and Clinical Psychology*, 84, 1–14. <https://doi.org/10.1037/a0039773>
- Kovacs, M. (2003). *Children's Depression Inventory (CDI): Technical manual update*. North Tonawanda, NY: Multi-Health Systems.
- Lock, S., & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change*, 20, 183–199. <https://doi.org/10.1375/bech.20.4.183.29383>
- Lowry-Webster, H. M., Barrett, P. M., & Lock, S. (2003). A universal prevention trial of anxiety symptomology during childhood: Results at 1-year follow-up. *Behaviour Change*, 20, 25–43. <https://doi.org/10.1375/bech.20.1.25.24843>
- Malmberg, M., Rydell, A. M., & Smedje, H. (2003). Validity of the Swedish version of the Strengths and Difficulties Questionnaire (SDQ-Swe). *Nordic Journal of Psychiatry*, 57, 357–363. <https://doi.org/10.1080/08039480310002697>
- McLeod, B. D., Sutherland, K. S., Martinez, R. G., Conroy, M. A., Snyder, P. A., & Southam-Gerow, M. A. (2017). Identifying common practice elements to improve social, emotional, and behavioral outcomes of young children in early childhood classrooms. *Prevention Science*, 18, 204–213. <https://doi.org/10.1007/s11211-016-0703-y>
- Merikangas, K. R., He, J. P., Brody, D., Fisher, P. W., Bourdon, K., & Koretz, D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001–2004 NHANES. *Pediatrics*, 125, 75–81. <https://doi.org/10.1542/peds.2008-2598>
- Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing risks for mental disorders: Frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Nauta, M. H., Scholing, A., Rapee, R. M., Abbott, M., Spence, S. H., & Waters, A. (2004). A parent-report measure of children's anxiety: Psychometric properties and comparison with child-report in a clinic and normal sample. *Behaviour Research and Therapy*, 42, 813–839. [https://doi.org/10.1016/S0005-7967\(03\)00200-6](https://doi.org/10.1016/S0005-7967(03)00200-6)
- Owens, J. S., Lyon, A. R., Brandt, N. E., Warner, C. M., Nadeem, E., Spiel, C., & Wagner, M. (2014). Implementation science in school mental health: Key constructs in a developing research agenda. *School Mental Health*, 6, 99–111. <https://doi.org/10.1007/s12310-013-9115-3>
- R Core Team (2015). R: A language and environment for statistical computing. Available at <https://www.R-project.org/>.
- Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., & O'Mahen, H. (2016). Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (COBRA): A randomised, controlled, non-inferiority trial. *The Lancet*, 388, 871–880. [https://doi.org/10.1016/S0140-6736\(16\)31140-0](https://doi.org/10.1016/S0140-6736(16)31140-0)
- Rooney, R. M., Morrison, D., Hassan, S., Kane, R., Roberts, C., & Mancini, V. (2013). Prevention of internalizing disorders in 9–10 year old children: Efficacy of the Aussie Optimism Positive Thinking Skills Program at 30-month follow-up. *Frontiers in Psychology*, 4, 1–10. <https://doi.org/10.3389/fpsyg.2013.00988>
- Rumberger, R. W. (2003). The causes and consequences of student mobility. *Journal of Negro Education*, 72, 6–21. <https://doi.org/10.2307/3211287>
- Sanchez, A. L., Cornacchio, D., Poznanski, B., Golik, A. M., Chou, T., & Comer, J. S. (2018). The effectiveness of school-based mental health services for elementary-aged children: A meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57, 153–165. <https://doi.org/10.1016/j.jaac.2017.11.022>
- Snell, T., Knapp, M., Healey, A., Guglani, S., Evans-Lacko, S., Fernandez, J. L., & Ford, T. (2013). Economic impact of childhood psychiatric disorder on public sector services in Britain: Estimates from national survey data. *Journal of Child Psychology and Psychiatry*, 54, 977–985. <https://doi.org/10.1111/jcpp.12055>
- Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, 36, 545–566. [https://doi.org/10.1016/S0005-7967\(98\)00034-5](https://doi.org/10.1016/S0005-7967(98)00034-5)
- Spence, S. H. (1999). *Spence Children's Anxiety Scale- Parent Version*. Brisbane, Australia: University of Queensland.
- Spence, S. H., Sheffield, J. K., & Donovan, C. L. (2005). Long-term outcome of a school-based, universal approach to prevention of depression in adolescents. *Journal of Consulting and Clinical Psychology*, 73, 160–167. <https://doi.org/10.1037/0022-006X.73.1.160>
- Stice, E., Shaw, H., Bohon, C., Marti, C. N., & Rohde, P. (2009). A meta-analytic review of depression prevention programs for children and adolescents: Factors that predict magnitude of intervention effects. *Journal of Consulting and Clinical Psychology*, 77, 486–503. <https://doi.org/10.1037/a0015168>
- Teubert, D., & Pinquart, M. (2011). A meta-analytic review on the prevention of symptoms of anxiety in children and adolescents. *Journal of Anxiety Disorders*, 25, 1046–1059. <https://doi.org/10.1016/j.janxdis.2011.07.001>
- Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, 24, 80–91. <https://doi.org/10.1016/j.tate.2007.01.004>

RECEIVED: September 15, 2017

ACCEPTED: May 14, 2018

AVAILABLE ONLINE: 19 May 2018