Evaluating the implementation of a school-based emotional well-being programme: a cluster randomized controlled trial of Zippy’s Friends for children in disadvantaged primary schools

Aleisha M. Clarke1*, Brendan Bunting2 and Margaret M. Barry1
1Health Promotion Research Centre, National University of Ireland Galway, Ireland and
2Psychology Research Institute, University of Ulster, Londonderry, UK
* Correspondence to: A. M. Clarke. E-mail: aleisha.clarke@nuigalway.ie
Received on December 6, 2013; accepted on July 15, 2014

Abstract
Schools are recognized as one of the most important settings for promoting social and emotional well-being among children and adolescents. This clustered randomized controlled trial evaluated Zippy’s Friends, an international school-based emotional well-being programme, with 766 children from designated disadvantaged schools. The purpose of this study was to evaluate the immediate and long term impact of the programme and to determine the impact of implementation fidelity on programme outcomes. Teachers reported emotional literacy outcomes using the Emotional Literacy Checklist, and emotional and behavioural outcomes using the Strengths and Difficulties Questionnaire. Controlling for the hierarchical structure of the data, path analysis using structural equation modelling revealed that the programme had a significant positive impact on the children’s emotional literacy scores including significant improvements in the subscale scores of self-awareness \(P < 0.001\), self-regulation \(P < 0.01\), motivation \(P < 0.001\) and social skills \(P < 0.001\) at post-intervention. These results were maintained at 12-month follow-up \(P < 0.01\). The programme, however, did not have a significant impact on children’s emotional and behavioural problems. Analysis of programme fidelity indicated that high fidelity was directly related to improved emotional literacy scores at post-intervention.

Introduction
It is widely recognized that in addition to fostering children’s academic development, schools play an important role in fulfilling a broader educational agenda that is concerned with enhancing students’ social-emotional well-being, character, health and civic engagement [1]. Evidence from systematic reviews shows that school-based mental health promotion interventions, when implemented effectively, can produce long-term benefits for young people, including improved emotional and social functioning and academic achievement [2–5]. Universal interventions that promote competence and life skills in the context of a whole-school approach have been found to produce significant positive outcomes for children’s mental health and well-being [4, 6–9].

The Zippy’s Friends programme is a universal school-based programme for children aged between 5 and 8 years. The programme is designed to promote the mental health and emotional well-being of all young children by increasing their repertoire of coping skills and by stimulating varied and flexible ways of coping with problems of day-to-day life [10]. Zippy’s Friends is similar to interventions...
such as ‘Al’s Pals: kids making healthy choices’ [11], ‘Promoting Alternative Thinking Strategies (PATHS)’ [12], and ‘I Can Problem Solve (ICPS)’ [13] in its focus on enhancing children’s social-emotional skills such as self-control, self-esteem, social skills, problem solving and decision making. Skill concepts are also presented in a similar format across these interventions, through interactive teaching methods including storytelling, discussion, modelling and role playing. The Zippy’s Friends programme builds on theory regarding the relationship between negative life events, coping and mental health [14, 15]. Previous research indicates that having a repertoire of coping strategies can help young children mitigate the effects of stressors on the development of psychological problems [16].

The Zippy’s Friends programme is shorter in comparison to PATHS and ICPS, which adopt a whole-school approach to implementation. Zippy’s Friends consists of 24 sessions implemented over one academic year. The 24 sessions are divided into six modules, each module containing four lessons which are conducted once a week for 1 h by the class teacher. Each module is centred around a set of six illustrated stories about a group of children, their families, friends and an imaginary stick insect called Zippy. The modules focus on a particular theme: (i) feelings, (ii) communication, (iii) making and breaking relationships, (iv) conflict resolution, (v) dealing with change and loss and (vi) general coping skills.

Zippy’s Friends is currently running in 27 countries and more than one million children in schools and kindergartens have enrolled in the programme worldwide. To date, the programme has been evaluated in Norway [17], Denmark and Lithuania [18] and Canada [19]. Key findings from these studies highlight the significant positive effects of the programme on children’s coping strategies [17, 18], emotional literacy skills [18], social skills [18], externalizing behaviour [18], improved autonomy [19] and reduced mental health difficulties in daily life [17]. Broader findings from the study conducted in Norway included improved social climate in the classroom, reduced bullying and improved academic skills [20].

The present study sought to build on previous programme evaluations by using a clustered randomized controlled trial to examine the immediate and long term impact of the programme on children in disadvantaged schools. Given that randomization occurred at school level and that children were nested within schools, individual student data were not independent. The study accounted for this nested group design. Several systematic reviews have drawn attention to the number of studies that fail to employ proper statistical procedures to account for nesting or clustering of data [3, 7, 9]. The study also aimed to investigate the impact of programme implementation on programme outcomes. Research on a wide range of mental health interventions has increasingly recognized the need to systematically evaluate the process of implementation and to determine its impact on programme outcomes [21, 22]. This article reports on (i) the impact of the programme on children’s emotional literacy skills, emotional and behavioural functioning, and (ii) the impact of programme fidelity on programme outcomes. Findings in relation to other aspects of the evaluation, including the effect of the programme on children’s coping skills using child participatory methods, and the assessment of other aspects of programme implementation (school context, quality of implementation, participant responsiveness) are reported elsewhere [23–25].

The programme was piloted in Ireland as part of the Social, Personal and Health Education (SPHE) curriculum. SPHE, which is mandatory in primary schools in Ireland, focuses on the development of a broad range of skills relevant to children’s physical, personal, emotional and social health within a supportive whole-school environment [26]. The evaluation was carried out in the context of ‘Delivering Equality of Opportunity in Schools’ (designated disadvantaged) schools in Ireland. Economically disadvantaged children are considered at risk for the development of social, emotional and behavioural problem because of the greater number of negative or undesirable life events and adverse conditions (risks factors) to which they are exposed [27–31]. Ethical approval was received from the National University of Ireland Galway Ethics
Committee in December 2007 to undertake the evaluation of the Zippy’s Friends programme.

**Methods**

**Study design and sample**

This study employed a cluster randomized controlled design, with assessments carried out before (T1), immediately after (T2) and at 12-month post-implementation (T3). The programme was implemented with children in first class (aged 7–8 years). To qualify for selection (i) schools had to be mixed gender school and assigned the designated disadvantaged status by the Department of Education and Skills and (ii) classes had to contain 10 or more children. A total of 48 schools fulfilled these criteria and 44 schools agreed to participate (participation rate 91.7%). Schools were randomly assigned intervention Type I, intervention Type II and control status.

Teachers in intervention Type I were asked to implement the programme as faithfully as possible. The teachers in intervention Type II were requested to use the programme as a resource. They could, therefore, decide which parts of the programme they would implement and could combine this programme with other SPHE curriculum resources. This was a requirement from the Department of Education and Skills in order to determine the effectiveness of the programme when utilized as a resource in comparison to full implementation of the programme. The control schools were given no direction and thus implemented the SPHE curriculum as usual. Table I outlines the number of schools, classes, teachers and children enrolled in study.

### Table I. Number of schools, classes, teachers and children enrolled in study

<table>
<thead>
<tr>
<th></th>
<th>Intervention Type I</th>
<th>Intervention Type II</th>
<th>Control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schools</td>
<td>15</td>
<td>15</td>
<td>14</td>
<td>44</td>
</tr>
<tr>
<td>Classes</td>
<td>18</td>
<td>18</td>
<td>16</td>
<td>52</td>
</tr>
<tr>
<td>Teachers</td>
<td>18</td>
<td>18</td>
<td>16</td>
<td>52</td>
</tr>
<tr>
<td>Children</td>
<td>267</td>
<td>277</td>
<td>222</td>
<td>766</td>
</tr>
</tbody>
</table>

**Context of implementation**

Implementation of the programme was coordinated by the Health Promotion Service of the Health Service Executive (HSE) in Ireland. A joint partnership between the HSE and the Department of Education and Skills was set up to prepare for programme implementation. Health Promotion Specialists provided a 2-day training workshop with intervention teachers. During programme implementation, the Health Promotion Specialists were also engaged in the provision of ongoing support to teachers through school visits and an interim group meeting.

**Measures**

**Children’s emotional literacy**

The teachers completed the Emotional Literacy Checklist [32] at pre-, post-intervention and 12-month follow-up. This questionnaire measures five dimensions of emotional literacy: self-awareness, self-regulation, motivation, empathy and social skills. All 20 items on the checklist are rated on a four-point Likert Scale. Sample items from this measure include, self-awareness ‘Can name his/her own feelings’; self-regulation ‘Is aware of his/her own strengths and qualities’; empathy ‘Is intolerant of people who are different from him/her’; motivation ‘Gives up easily when faced with something difficult’; social skills ‘laughs and smiles when it is appropriate to do so’. The Emotional Literacy Checklist has good internal consistency with a Cronbach’s alpha coefficient reported of 0.94 [32]. In this study, the Cronbach alpha coefficient was 0.91.

**Children’s emotional and behavioural functioning**

The teachers also completed the Strengths and Difficulties Questionnaire [33]. This standardized
questionnaire measures children’s (age 4–16 years) emotional and behavioural functioning. The 25-item questionnaire generates five main subscale scores: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behaviour. According to Goodman [34], the Strengths and Difficulties Questionnaire has good internal consistency with a Cronbach’s alpha coefficient of 0.73. The Cronbach’s alpha coefficient in this study was 0.76.

Programme fidelity
After each Zippy’s Friends session, teachers completed a programme fidelity checklist indicating what parts of each session were fully implemented, partially implemented and omitted. This checklist was part of the teachers’ weekly questionnaire, which was designed to elicit information about the implementation process. Programme fidelity was also monitored as part of structured class observations which were carried out by the researcher (A.M.C) and a Health Promotion Specialist in a sample of classes over the course of the study (N = 27). After observing the lesson, the researcher/Health Promotion Specialist completed an observation questionnaire which included the same fidelity checklist that teachers were required to complete after each session. The Kappa Measure of Agreement value between the two raters was 0.738 with a significance value of P < 0.005.

Data analysis
Because of the hierarchical structure of the data, path analyses using structural equation modelling were conducted using Mplus (Version 7.10), controlling for the nested data structure (individuals were nested within classes). In this study, schools were randomized to one of two conditions. Randomization occurred at this level in order to minimize the effects of cross contamination of the intervention, and because of the class/grade selection criterion in place (children in first class), this meant that most schools only contributed one class from within the school, consequently, there was little difference in the design effects for these levels. The analysis was conducted in terms of the occasions and at the level of the individual, because it was at these levels that the intervention was expected to have an effect. The model was expanded to include analysis of covariance within the structural equation modelling framework in order to correct for measurement error and adjusting for the imbalance in scores across the intervention and control group at the baseline. Relevant factor loadings and intercepts were restricted to be equal.

Model fit was evaluated using a chi-square test statistic as well as Comparative Fit Index (CFI), Tucker-Lewis Index (TLI) and root-mean-square error of approximation (RMSEA). A relatively good model fit is described by values of CFI and TLI greater than 0.95 and RMSEA values less than 0.06 which indicated the absence of large differences between the models and the data [35, 36]. Analyses were conducted with an adjustment for non-normality using full-information maximum-likelihood estimation, which is considered a viable strategy for addressing missing data [37, 38]. Because the latent variable mean scores for categorical data in Mplus are standardized to zero, descriptive values (mean and standard deviation) were obtained from analysis in SPSS Version 21. Standardized values are reported as an index of effect size. The standardized values can be interpreted as an effect size for a 1U change (standard deviation) in an exogenous measure (independent) on an endogenous measure (dependent).

Results
Participant profile
Table II presents information on the gender profile of the participants, the nature of the classes and school location. The schools varied in size and location (urban/rural), however, there was no significant difference between the intervention and control groups in terms of gender, school location (rural or urban) and multigrade class (multiple classes in one class unit or single class unit). The mean age of the
children across the three groups (7 years, 3 months) did not differ significantly.

The flow chart in Fig. 1 tracks the loss of children and classes over the course of the study. The main reasons for attrition included schools no longer being able to commit to the study, teachers not returning questionnaires and children moving to another school.

Programme fidelity

The mean number of activities that the teachers in intervention Type I and Type II reported implementing in full, in part and omitting are presented in Table III. The results indicate that there was no significant difference in programme fidelity between the two intervention groups. Programme fidelity was high among both groups. These results are supported by fidelity findings from the structured observations ($N = 27$) carried out in a sample of classes over the course of the study (observations: mean number of observed activities full implementation = 90.9%; partial implementation = 2.3%; activities not implemented = 6.8%) Given that there was no significant difference in the level of programme adherence between the intervention groups, the impact results were analysed by comparing the intervention group’s results (i.e. intervention Type I combined with intervention Type II) with the control group’s results.

Children’s emotional literacy

The mean values for the intervention and control groups’ subscale scores at pre-, post-intervention and 12-month follow-up are presented in Table IV. Baseline differences were controlled for using analysis of covariance within the structural equation modelling framework.

To address the first hypothesis and investigate the effect of the Zippy’s Friends intervention on the children’s emotional literacy skills at post-intervention and at 12-month follow-up, a model for each subscale was fitted. Figure 2 illustrates the theoretical model for the subscale self-awareness. Each model consisted of four items at three points in time (pre-intervention, post-intervention and 12-month follow-up) and a group variable (intervention/control group). Correlated errors for each item at pre, post and 12-month follow-up were fitted. All standardized factor loadings were greater than 0.7. The same theoretical model was tested for each of the subscales, estimating (i) the direct effect of the intervention at post-intervention and (ii) the indirect effect of the post-intervention score on the 12-month follow-up score.

Table V presents the model fit statistics (chi square, CFI, TLI and RMSEA) for all subscales. All of the models provided an adequate fit to the data with relatively low values of chi-square ($P < 0.05$) and relatively high values of CFI and TLI indices. The values of RMSEA were less than 0.06. Inter-correlation analysis was conducted at the baseline to determine the level of correlation between the subscales. All correlations were less than 0.5 with the exception of empathy and self-regulation ($r = 0.88$), motivation and self-awareness ($r = 0.74$), social skills and self-awareness ($r = 0.77$), empathy and social awareness ($r = 0.6$).

Post-intervention results revealed that there was a significant direct effect of the programme on the intervention group’s self-awareness [(Estimate $= 0.39$, SE $= 0.057$, C.R. $= 6.875$, $P < 0.001$), Std Est $= 0.351$]; self-regulation [(Estimate $= 0.220$, SE $= 0.083$, C.R. $= 2.66$, $P < 0.01$); Std Est $= 0.122$]; motivation [(Estimate $= 0.215$, SE $= 0.058$,
C.R. = 3.691, \( P < 0.001 \), Std Est = 0.133] and Social Skills score [(Estimate = 0.215, SE = 0.058, C.R. = 3.691, \( P < 0.001 \), Std Est = 0.124]. These results indicate a significant increase in the intervention group’s emotional literacy skills across these four subscales between pre- and post-intervention. There was no significant direct programme effect on the intervention group’s Empathy score at post-intervention [(Estimate = 0.072, SE = 0.060, C.R. = 1.203, \( P = 0.229 \), Std Est = 0.065].

Twelve-month follow-up results revealed a similar pattern. There was an additional indirect programme effect on the intervention group’s self-awareness [(Estimate = 0.155, SE = 0.049, C.R. = 3.186, \( P < 0.01 \), Std Est = 0.142]; self-regulation [(Estimate = 0.107, SE = 0.048, C.R. = 2.211, \( P < 0.05 \), Std Est = 0.059]; motivation [(Estimate = 0.094, SE = 0.036, C.R. = 2.587, \( P < 0.01 \), Std Est = 0.054] and Social Skills scores [(Estimate = 0.094, SE = 0.036, C.R. = 2.587,
Empathy had no immediate or delayed impact on the subscale EST \(= SE \times C.R. \times P \times 0.034, C.R. = 1.047, P = 0.295\), Std Est = 0.028], thus indicating that the programme had no immediate or delayed impact on the subscale Empathy.

### Children’s emotional and behavioural functioning

The mean values of the intervention and control groups’ Strengths and Difficulties (Goodman, 1997) subscales at pre-, post-intervention and follow-up are presented in Table VI. Baseline differences for the subscales emotional symptoms, hyperactivity, peer relationship problems and prosocial behaviour were controlled for using analysis of covariance. Mean scores indicate an improvement in intervention and control group subscales scores between pre- and post-intervention. Inter-correlation analysis revealed that correlations were less than 0.5 with the exception of conduct problems and hyperactivity \((r = 0.69)\) and peer relationship problems and conduct problems \((r = 0.58)\).

To investigate the effect of the Zippy’s Friends intervention on the children’s emotional and behavioural functioning at post-intervention and at 12-month follow-up, a model for each subscale (emotional symptoms, hyperactivity, conduct problems, peer relationship problems and prosocial behaviour) was fitted. Table V, which presents the model fit statistics (chi square, CFI and RMSEA) for all subscales, indicates that all of the models provided an adequate fit to the data. Post-intervention analyses revealed that the intervention had no significant effect on the intervention group’s emotional symptoms \([(Estimate = -0.068, SE = 0.51, C.R. = -1.32, P = 0.188), Std, Est = -0.109]; hyperactivity \([(Estimate = -0.046, SE = 0.077, C.R. = -0.604, P = 0.546), Std, Est = -0.037]; peer relationship problems \([(Estimate = -0.054, SE = 0.047, C.R. = -1.141, P = 0.254), Std, Est = 0.094]; and prosocial behaviour \([(Estimate = 0.058, SE = 0.054, C.R. = 1.075, P = 0.282), Std, Est = 0.076]. Regarding the subscale conduct problems, children in the control group evidenced a significant decrease in their score between pre- and post-intervention when compared with the intervention group \([(Estimate = -0.132, SE = 0.059, C.R. = -2.242.003, P < 0.05), Std, Est = -0.141]. Twelve-month follow-up results revealed that the improvement in the control group’s conduct problems was not sustained at 12-month follow-up \([(Estimate = -0.079, SE = 0.057, C.R. = -1.386, P = 0.166), Std, Est = -0.078].

Direct programme effects at 12-month follow-up were tested on subscales that showed no programme effect at post-intervention in order to determine if there was a delayed programme effect. Analyses revealed no direct programme effect at 12-month follow-up across the four subscales: emotional symptoms \([(Estimate = -0.041, SE = 0.4, C.R. = -1.043, P = 0.297), Std, Est = -0.061]; hyperactivity \([(Estimate = -0.075, SE = 0.08, C.R. = 0.936, P = 0.349), Std, Est = -0.06]; peer

| Table IV. Descriptive statistics for Emotional Literacy Checklist subscales |
|-----------------|-----------------|-----------------|
|                  | Pre-intervention | Post-intervention | Follow-up |
|                  | Mean | SD   | Mean | SD   | Mean | SD   |
| Self-awareness   |      |      |      |      |      |      |
| Interv           | 11.53 | 2.1  | 13.04 | 2.1  | 12.47 | 1.9  |
| Control          | 12.24 | 2.3  | 12.92 | 1.8  | 12.69 | 2.0  |
| Self-regulation  |      |      |      |      |      |      |
| Interv           | 11.69 | 3.2  | 12.54 | 2.9  | 12.48 | 2.9  |
| Control          | 12.05 | 3.2  | 12.33 | 2.8  | 12.89 | 2.8  |
| Motivation       |      |      |      |      |      |      |
| Interv           | 11.48 | 3.3  | 12.37 | 2.8  | 12.03 | 2.9  |
| Control          | 12.13 | 3.0  | 12.07 | 2.6  | 12.09 | 3.1  |
| Empathy          |      |      |      |      |      |      |
| Interv           | 12.76 | 2.6  | 13.42 | 2.3  | 13.19 | 2.4  |
| Control          | 13.15 | 2.4  | 13.57 | 2.0  | 13.65 | 2.3  |
| Social skills    |      |      |      |      |      |      |
| Interv           | 13.50 | 2.3  | 14.30 | 1.8  | 14.13 | 1.8  |
| Control          | 13.94 | 2.1  | 14.35 | 1.6  | 14.80 | 1.4  |

Mean scores do not account for baseline differences controlled for in analyses.
relationship problems [(Estimate = −0.024, SE = 0.068, C.R. = −0.347, P = 0.729), Std, Est = −0.035]; and prosocial behaviour [(Estimate = 0.061, SE = 0.067, C.R. = 0.915, P = 0.37), Std, Est = 0.075]. These results indicate that the intervention group’s emotional and behavioural problems did not improve as a result of the programme.

**Programme fidelity**

In order to determine the impact of fidelity on programme outcomes, a separate path analysis was
Table VI. Descriptive statistics for Strengths and Difficulties subscales

<table>
<thead>
<tr>
<th></th>
<th>Pre-intervention</th>
<th>Post-intervention</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean* SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
</tr>
<tr>
<td>Emotional symptoms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interv</td>
<td>2.42 2.5</td>
<td>1.68 1.9</td>
<td>2.07 2.1</td>
</tr>
<tr>
<td>Control</td>
<td>1.89 2.0</td>
<td>1.52 1.5</td>
<td>1.65 1.7</td>
</tr>
<tr>
<td>Conduct problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interv</td>
<td>1.51 2.0</td>
<td>1.22 1.6</td>
<td>1.25 1.7</td>
</tr>
<tr>
<td>Control</td>
<td>1.59 2.1</td>
<td>1.07 1.5</td>
<td>1.11 1.6</td>
</tr>
<tr>
<td>Hyperactivity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interv</td>
<td>3.77 3.1</td>
<td>3.24 2.9</td>
<td>3.15 2.7</td>
</tr>
<tr>
<td>Control</td>
<td>3.48 3.2</td>
<td>3.28 2.5</td>
<td>2.95 2.7</td>
</tr>
<tr>
<td>Peer relationship problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interv</td>
<td>1.66 1.9</td>
<td>1.15 1.4</td>
<td>1.26 1.5</td>
</tr>
<tr>
<td>Control</td>
<td>1.74 2.2</td>
<td>1.17 1.3</td>
<td>1.37 1.7</td>
</tr>
<tr>
<td>Prosocial behaviour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interv</td>
<td>7.04 2.3</td>
<td>7.68 2.0</td>
<td>7.70 2.1</td>
</tr>
<tr>
<td>Control</td>
<td>7.57 2.3</td>
<td>7.97 1.9</td>
<td>8.31 1.7</td>
</tr>
</tbody>
</table>

*Mean scores do not account for baseline differences controlled for in analyses.

conducted with the intervention group’s fidelity scores and emotional literacy subscales that improved significantly between pre- and post-interventions. Each subscale model consisted of the subscale items at two time points (pre- and post-intervention) and a fidelity variable based on the teacher’s level of fidelity (expressed as a percentage of total activities fully implemented across the 24 sessions). Model fit was again evaluated by examining chi square test statistic, CFI, TLI and RMSEA values. All of the models provided an adequate fit to the data.

Path analysis revealed that after controlling for clustering and all factors restricted to being equal across the time points, programme fidelity had a significant impact on the Emotional Literacy subscale scores at post-intervention: self-awareness: [(Estimate = 0.025, SE = 0.006, C.R. = 4.38, P < 0.001), Std Eff = 0.087]; self-regulation: [(Estimate = 0.024, SE = 0.006, C.R. = 4.04, P < 0.001), Std Eff = 0.08]; motivation: [(Estimate = 0.022, SE = 0.006, C.R. = 3.73, P < 0.001), Std Eff = 0.075] and social skills [(Estimate = 0.022, SE = 0.005, C.R. = 4.02, P < 0.001), Std Eff = 0.081]. These results indicate that fidelity had a direct impact on the intervention group’s emotional literacy scores at post-intervention, in that, the higher the level of fidelity, the higher the children’s emotional literacy subscale scores at post-intervention.

### Discussion

This clustered randomized controlled trial examined the immediate and long term impact of the Zippy’s Friends programme on children in disadvantaged schools and assessed the impact of programme fidelity on programme outcomes. Path analysis revealed that the programme had a significant positive effect on the children’s emotional literacy skills. Post-intervention results from the Emotional Literacy Checklist [32] showed a significant increase in the intervention group’s subscale scores on self-awareness, self-regulation, motivation and social skills when compared with the control group. These positive results were supported by additional qualitative findings from the teachers’ end of programme review questionnaire and focus group sessions which were also carried out at post-intervention [23–25]. Twelve-month follow-up findings point to the significant lasting effect of the programme on the children’s emotional literacy skills, with the exception of one subscale, empathy. The long-term programme effect may be due to the fact that Zippy’s Friends materials provided structure, consistency and repetition in the delivery of content in relation to emotional literacy skills. Qualitative results from this study also underscore the importance of these practices in facilitating programme implementation and achieving positive programme outcomes from the teachers’ perspectives [24–25]. These findings are in keeping with evidence from a meta-analysis by Durlak et al. [3], which showed larger effect sizes (0.69 versus 0.01) for school-based social-emotional interventions (N=213 interventions involving 270,034 children aged 5–13 years) that followed principles including, structure, sequenced learning, active involvement and repetition. In this study, however, it is important to note that with the exception of the intervention
effect on self-awareness ($\beta = 0.35$), statistically significant intervention effects were small ($<0.15$). A possible reason for this is the nature of the intervention in that it is a universal programme as opposed to a selective intervention targeting children at higher risk of emotional behavioural problems. Baseline results from the Emotional Literacy Checklist indicate relatively high emotional literacy scores across the intervention and control group suggesting a possible ceiling effect at post-intervention and follow-up. Previous reviews have identified the issue of the ‘ceiling effect’, with populations that do not have overt problems not having the same scope for improvement [4, 7].

Despite the programme’s positive impact on children’s emotional literacy skills, results from the Strengths and Difficulties Questionnaire [33] revealed that the programme did not have a significant positive effect on the intervention group’s emotional and behavioural problems including the subscales, emotional symptoms, hyperactivity, peer relationship problems and prosocial behaviour. Similar results have been reported in an evaluation of Zippy’s Friends in Norway which found that neither the teachers nor the parents reported improvements in children’s emotional and behavioural problems as measured by the Strengths and Difficulties Questionnaire [17]. An unexpected finding from the Strengths and Difficulties Questionnaire was the significant improvement in the control group’s Conduct Problems subscale score between pre- and post-intervention. Whilst it is difficult to explain this end of programme finding, it is important to note that this result was not sustained at 12-month follow-up thus indicating that there was no long-term improvement in the control group’s conduct problems when compared with the intervention group.

A likely explanation for the lack of immediate and long-term findings on the intervention group’s emotional and behavioural symptoms may be due to the fact that training and support in classroom management and dealing with behavioural problems is not provided explicitly as part of the Zippy’s Friends programme. The inclusion of such training, as in other programmes such as ‘PATHS’ [12] could possibly improve children’s outcomes in relation to emotional and behavioural problems. Alternatively, the relatively short duration of the programme could be an alternative explanation, endorsing previous findings that a more integrated model with a sequenced curriculum as part of a whole-school approach would lead to more enduring behavioural and emotional outcomes [4, 7]. Greenberg et al. [39] argue that ‘short-term prevention interventions produce time-limited benefits at best with at-risk groups, whereas multi-year programmes are more likely to foster enduring benefits’ (p37). In the case of the Zippy’s Friends programme, the findings from this study suggest the need for an integrated model with a sequenced curriculum implemented through the primary school grades in combination with strategies promoting positive mental health at the school environment and community level.

Programme fidelity results indicate that the level of programme adherence across the two intervention groups was high (>86%). The fact that teachers in intervention Type II, who were given the option to implement the programme as a resource, implemented the programme with high fidelity is indicative of the level of teacher commitment and support for the programme. These data also point to the likelihood of future faithful implementation of the programme when rolled out more broadly. Analysis of the impact of programme fidelity on changes in children’s emotional literacy scores revealed that high fidelity was directly related to improved emotional literacy scores at post-intervention. These findings are consistent with a recent large-scale evaluation of ‘KidsMatter’, the primary school mental health initiative in Australia which found that children’s ($N = 4970$) social and emotional competencies significantly improved across average-to-high implementing schools but not in low-implementing schools [40]. The results from the Zippy’s Friends study underscore the need for high fidelity in achieving programme outcomes. Ultimately, these findings point to the role of programme implementation in the effectiveness of school-based interventions and highlight the need for programme evaluators to afford a higher priority to monitoring and reporting programme fidelity.
Limitations
In considering the findings of the present study, it is important to discuss its limitation. First, although the schools were randomly assigned to intervention and control groups, they were not randomly selected to participate in the study. All schools volunteered to participate in the trial and they may, therefore, be more typical in terms of schools that value the role of mental health promotion in education. Given this potentially applied to all schools, however, this cannot explain the difference in programme outcomes between the intervention and control groups. Related to the measures used, a limitation was the use of teacher self-report measures to evaluate programme effectiveness. In order to strengthen the reliability of the outcome findings, parent report questionnaires would have been a useful addition to the measures used in this study. Another limitation regarding measures is the high level of inter-correlation across some of the Emotional Literacy Checklist subscales. However, it is important to note that the subscales empathy and self-regulation, which were the most highly correlated ($r = 0.88$), differed in terms of programme outcomes. Whilst the intervention had a significant positive effect on children’s self-regulation at post-intervention and 12-month follow-up, there was no programme effect on children’s Empathy scores. This finding indicates that even when subscales were highly correlated, the intervention had a different effect on them, which would, therefore, suggest that these subscales are measuring two different constructs. In terms of measuring programme fidelity, the use of self-report questionnaires may have been coloured by socially desirable responses or self-reflective blind spots. Class observations, which were carried out in a sample of interventions schools, were used to address this issue and check for reliability of responses. In addition, because the participating schools were recruited from target areas where support could be provided by Health Promotion Specialists in piloting the programme, the results may not be representative of schools across the country as a whole. Finally, the programme was implemented with children attending disadvantaged schools, which also needs to be considered in interpreting the findings at a broader level.

Conclusions
Overall, the findings from this study indicate that teachers’ implementation of a universal emotional well-being programme with high fidelity resulted in significant gains for the children in terms of improved emotional literacy skills at post-intervention and at 12-month follow-up. The fact that the programme did not have a long term impact on the children’s emotional and behavioural functioning points to the need for additional resources in addressing these problems as part of a more sustained whole-school approach. A strength of the current study is the examination of programme fidelity and its impact on programme outcomes, which few studies have examined to date. At a broader level, the findings from this study contribute to the advancement of knowledge regarding the evaluation, implementation and effectiveness of school-based mental health promotion interventions in the primary school setting.

Acknowledgements
The authors are grateful to the staff and pupils of all the schools who participated in the study and to the Health Promotion Specialists and members of the National Advisory Board for their support and contribution to the research. The views expressed in this article are solely those of the authors.

Funding
HSE Population Health, Health Promotion Directorate; National Office for Suicide Prevention Ireland and the Irish Research Council for the Humanities and Social Sciences.

Conflict of interest statement
None declared.
References


