The effectiveness and micro-costing analysis of a universal, school-based, social-emotional learning programme in the UK: A cluster-randomised controlled trial

Berry, Vashti [a]
Axford, Nick [b]
Blower, Sarah [c]
Taylor, Rod S [d]
Edwards, Rhiannon Tudor [e]
Tobin, Kate [b]
Jones, Carys [e]
Bywater, Tracey [c]

[a] NIHR CLAHRC South West Peninsula (PenCLAHRC) Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK
[b] Dartington Social Research Unit, Dartington, UK
[c] Department of Health Sciences, University of York, York, UK
[d] Institute of Health Research, University of Exeter Medical School, Exeter, UK
[e] Centre for Health Economics and Medicines Evaluation, Bangor University, Bangor, UK

Corresponding author:
Vashti Berry, Room N10, ITTC Building, Plymouth Science Park, Plymouth, Devon, PL4 8AA, UK. Email: vashti.berry@plymouth.ac.uk

This is the accepted version of this article. The final version of this article was published online in the journal *School Mental Health* on 28th October 2015. The final publication is available at Springer via http://dx.doi.org/DOI 10.1007/s12310-015-9160-1

Abstract
There are a growing number of school-based interventions designed to promote children's social and emotional learning. One such intervention, PATHS (Promoting Alternative Thinking Strategies), was evaluated in a randomised controlled trial involving 5,074 pupils aged 4-6 years at baseline in 56 primary schools across a large city in the UK. The programme was implemented for two academic years. The primary outcome measure was the teacher-rated Strengths and Difficulties Questionnaire (SDQ). A secondary measure was the PATHS Teacher Rating Scale (PTRS). Observations of child and teacher behaviours were undertaken in a third of intervention and control schools using the Teacher-Pupil Observation Tool (T-POT). Regarding fidelity, dose and adherence were measured via weekly logs completed by teachers, and a semi-structured questionnaire completed by PATHS coaches was used as a global measure of fidelity (capturing adherence, dose and quality). A cost-consequence analysis examined programme costs from a multi-agency public sector perspective. At one year post-baseline there were no statistically significant differences between the programme and control groups on the SDQ subscales or the SDQ total difficulties and impact scores. There were statistically significant differences favouring the programme group for six out of 11 sub-scales on the secondary outcome measure (PTRS). At two years post-baseline, there were no statistically significant differences between the groups on either measure. Fidelity, according to the global measure, was relatively strong, and there was no relationship between fidelity and treatment effects. The average cost of PATHS was £12,666 per school or £139 per child. The study, which was fully powered and independent of the programme developer, shows no statistically significant effect of the programme on child behaviour or emotional well-being.

Introduction

Promoting Alternative Thinking Strategies (PATHS) is a school-wide programme designed to impact on the social and emotional learning (SEL) of children aged 4-11 years. Initially developed for children with hearing disabilities, it has been subjected to several trials in the US where it is considered to be a 'Model' evidence-based programme (Blueprints, 2012; Greenberg et al., 2002). A variation of the programme has been tested on US pre-kindergarten children (i.e. 4-5 year-olds) (Domitrovich et al., 2007).

PATHS is one of a growing number of interventions designed to promote children's SEL, defined as 'the processes through which children and adults acquire and effectively apply the knowledge, attitudes and skills necessary to understand and manage emotions, set and achieve positive goals, feel and show empathy for others, establish and maintain positive relationships, and make responsible decisions' (Collaborative for Academic, Social and Emotional Learning (CASEL), 2014). The skills are thus *intrapersonal*, such as being able to manage one's emotions, and *interpersonal*, such as establishing positive relationships (Humphrey, 2013).

Interventions to promote SEL vary considerably. A taxonomy developed by Humphrey (2013) suggests that they may: (i) be universal or targeted; (ii) focus on delivering taught curriculum, modifying the school environment, working with parents and the wider community, or a combination of these elements; and (iii) involve detailed structured guidance ('top-down') or emphasise flexibility and local adaptation ('bottom-up'). Typically SEL

interventions cover emotional competence skills (e.g. self-awareness, self-regulation, social awareness) and relational/pro-social skills (e.g. relationship skills, responsible decision-making) (Denham & Brown, 2010). These skills develop over time (Weare, 2004), so interventions need to be developmentally aligned and sequenced. Critiques of the SEL movement focus on the implication that childhood is in crisis, the assumed need for explicit instruction and the alleged cultural insensitivity (e.g. Ecclestone & Hayes, 2008; Furedi, 2008; Hoffman, 2009) but there are strong counter-arguments to some of these and efforts to address others (see Humphrey, 2013).

Many of these programmes are listed on the website of CASEL (Collaborative for Academic, Social and Emotional Learning).[1] A systematic review found that universal school-based SEL programmes have direct positive impacts on children's emotional and behavioural health, with effect sizes ranging from 0.2 to 0.6, and produce indirect effects on educational achievement (an 11-percentile points gain), which were explained by better-behaved, happier pupils learning more (Durlak et al., 2011).

PATHS is one of the better-known evidence-based SEL programmes. It is targeted at children aged 4-11 years and comprises separate volumes of lessons for each year group, all with developmentally appropriate pictures, photographs and posters. In each year level PATHS covers self-control, emotional understanding, positive self-esteem, relationships, and

[1] www.casel.org

interpersonal problem-solving skills. Each unit builds hierarchically upon and synthesises previous learning. The programme is designed to be taught by regular teachers two to three times per week, ideally in 20-30 minute lessons, with daily activities to promote generalisation. Lesson activities include dialoguing, role-playing, story-telling by teachers and peers, social and self-reinforcement, attribution training and verbal mediation. Pupils are helped to relate cognitive-affective understanding to real-life situations. The manual contains detailed lesson plans and scripts but there is flexibility so that individual teachers can adapt them to fit their own style and events in the classroom or school. Parent letters and home activity assignments are provided to encourage parent involvement and support.

There have been numerous studies of the impact of PATHS (in some cases PATHS was combined with another intervention), ranging in sample size from 79 children to over 5,000, and including eight randomised controlled trials (RCTs), of which the RCT reported in this paper is the largest. Most studies have been with children aged 5-11 years, including those with special educational needs, but some have involved pre-kindergarten children (4-5 year-olds). Although studies were mostly conducted in the US (Greenberg & Kusché, 1998; CPPRG, 1999, 2010; Kam et al., 2004; Seifer et al., 2004; Riggs et al., 2006; Domitrovich et al., 2007; Bierman et al., 2008), trials have also been completed in Northern Ireland (Ross et al., 2012), Switzerland (Malti et al., 2011) and now England (reported here). An additional, ongoing, trial in Manchester, UK, is testing the effects of the primary school programme

as children transition into secondary school.[2]

Across PATHS studies, outcome measures have focused mainly on emotion recognition, behaviour and social problem-solving, and included a mixture of observation and reports by teachers, parents and children. In the US the results have generally been positive. In summary, and relative to control groups, they have shown: lower rates of conduct problems, aggression, anxiety, depression, anger and Attention Deficit and Hyperactivity Disorder (ADHD) symptoms; and higher rates of social problem-solving (including peer conflicts), emotional understanding (recognising one's own and other people's feelings), self-control, peer sociability, school functioning, empathy for others, authority acceptance, cognitive concentration and social competence (Blueprints, 2014). However, the RCT in Switzerland failed to demonstrate the expected impact of PATHS on overall child outcomes at post-test, although it was effective for high-risk children (Malti et al., 2011). In Northern Ireland, significant, but only small, effects were demonstrated for the PATHS programme in the first year of implementation (Ross et al., 2012).

PATHS was selected for implementation and testing in a large city in the UK as part of a broader strategy called 'Brighter Futures' (Little et al., 2012). The programme was intended to enhance social-emotional wellbeing, and thus help prevent, or reduce, behaviour problems and conduct disorder. Behaviour problems are characterised by aggression and defiance as well as negative affect and deficits in peer relationships and prosocial behaviours. Conduct

[2] For the trial protocol see http://www.nets.nihr.ac.uk/__data/assets/pdf_file/0019/81712/PRO-10-3006-01.pdf.

disorders meet the diagnostic criteria captured in the ICD-10 (World Health Organization, 2004) or DSM-5 (American Psychiatric Association, 2013) and are marked by oppositionality, defiance and aggression.

A city-wide survey involving 500 parents of children aged 0-6, using the parent version of the Strengths and Difficulties Questionnaire (SDQ: Goodman, 1997), found a significantly greater proportion of children aged 5-6 years (n=80) were at risk of a conduct disorder (19%) compared to 7% in Great Britain as a whole (Hobbs et al., 2011). A parallel survey, involving over 10,000 young people aged 7-18, and using the SDQ self-report measure, found that a significantly greater number of children aged 11-15 in the city (n=3,293) were likely to meet a clinical diagnosis for conduct disorder (21%) compared with Great Britain as a whole (11%) (Hobbs et al., 2011). A range of evidence-based programmes at each stage of children's development were selected to address these and other difficulties, and tested by RCT (Little et al., 2012). PATHS was introduced as a universal prevention programme in primary schools.

It was considered important to evaluate PATHS by RCT in the UK because nearly all impact evaluations of SEL programmes have been US-based (see Durlak et al., 2011). The programme developers at the Prevention Research Centre at Penn State University supported the implementation of PATHS in the UK city but were not involved in the evaluation. The study reported here is therefore one of the largest independent tests of an established SEL programme in Europe, the most robust test to date of PATHS in Great Britain

– others have been non-experimental (e.g. Curtis & Norgate, 2007) – and one of the few around the world undertaken as part of a public system's strategy to improve child outcomes.

In recent years there have been several attempts to introduce SEL programmes into schools in England. Early in 2003, a review for the UK Department for Education recommended that schools develop and adopt programmes designed to promote social emotional competence (Weare & Gray, 2003). The SEAL (Social and Emotional Aspects of Learning) programme was introduced by the then government (DfES 2005, 2007) and by 2010 was used in an estimated 90% of primary schools and 70% of secondary schools nationally (Humphrey, 2013). (For background to the SEAL programme see Bywater & Sharples, 2012.) It focuses on the skills of social awareness, self-regulation, motivation, empathy and social skills, and combines a taught curriculum with school-wide activities and inputs for teachers. It might be considered an amalgam of existing programmes on the CASEL list, drawing on the expertise of English educationalists. The national evaluation of SEAL showed mixed effects on outcomes in primary school, and no impact at secondary school (Humphrey et al. 2008, 2010).

Since SEAL was widely used in the city concerned before and to a lesser extent *during* the PATHS trial it is important to note how the two compare. Although both programmes are universal, serving all children rather than a targeted sub-group, and cover similar ground in terms of themes, there are two main differences. The first is that whereas PATHS is primarily a single-

component approach, with a strong focus on teaching an explicit curriculum, SEAL is a multi-component approach, with a greater emphasis on creating a positive school climate and ethos and providing continuing professional development for school staff (Humphrey, 2013). Second, SEAL is more bottom-up compared with the more top-down approach of PATHS: it represents a 'loose enabling framework' rather than a structured 'package' (Weare, 2010). As such, it might be considered 'the diametric opposite of the top-down manualised SEL interventions that have been popular in the US' (Humphrey, 2013: 57) – of which PATHS is emblematic.

This paper presents a cluster RCT evaluation of PATHS involving 5,074 children in 183 classrooms (across Reception, Year 1 and Year 2 grades[3]), in 56 primary schools in a large city in the UK. The study sought to examine the effectiveness and cost-effectiveness of PATHS for reducing children's level of behavioural and emotional difficulty. It was considered important to measure the cost and cost-effectiveness of the intervention given the climate of public sector austerity and in order to inform potential scale-up.

Method

Design

This study was a pragmatic, cluster trial with random allocation of schools on a 1:1 basis to receive the intervention or to a waiting list control, stratified by size of school and percentage of children qualifying for free-school meals (an

[3] This concerns children aged 4-7, equivalent to Pre-kindergarten, Kindergarten and 1st Grade in the US.

indicator of social disadvantage). It followed children within classrooms, within schools, over two years, collecting data at three points in time (baseline, 12 months post-baseline and 24 months post-baseline). The study was funded by Birmingham City Council, and received ethical approval from an independent research ethics committee (The Warren House Ethics Committee, 2009).

Participants

All mainstream (i.e. not special schools) primary schools in Birmingham (n = 299) were invited to take part in the study. They were provided with information about the PATHS programme, the intended RCT evaluation, and the expectations of participating schools. Head teachers committing to take part in the study were asked to sign a consent form indicating that they understood the above. A total of 64 schools expressed an initial commitment to taking part. This relatively low return (21%) may reflect the fact that SEAL was already operating in the city, leading arguably to a mix of scepticism and inertia regarding another SEL programme. All of the head teachers who agreed for their schools to participate in the study indicated a willingness to deliver the PATHS programme should their school be allocated to the intervention group. Participants were boys and girls in Reception and Year One grades of the participating schools in the 2009/2010 academic year who went on to Year One and Year Two, respectively, in 2010/2011. In the UK, this equates to children aged from 4-5 years to children aged 6-7 years. The PATHS programme can operate across all primary school years on a rolling basis but in this case it was decided to focus initially at least on the younger

children. There were no exclusion criteria for mainstream primary schools. All schools (intervention and control) had access to the national SEAL curriculum in the years prior to this pragmatic evaluation of the PATHS programme.

Sample size

Based on a large body of existing experimental evidence on the PATHS curriculum, and a review of other school-based SEL programmes, an effect size estimate for the intervention was judged to be 0.2 to 0.3 (i.e. a change of 20 to 30 percent of a standard deviation) improvement in social and emotional well-being (see especially CPPRG 1999 and Durlak et al., 2011). Software developed by Stephen Raudenbush and colleagues was used to calculate ideal and sufficient recommendations for the number of schools needed in the study (see http://sitemaker.umich.edu/group-based/optimal_design_software). It was estimated that a sample of 60 schools was needed, assuming an average cluster size of 90 children, an assumed Intra-Class Correlation (ICC) of 0.15 and a covariate of 0.20, to detect an effect size of 0.23 with power at 80% at an alpha of 5% or 50 schools to detect an effect size of 0.30 at 80% power and alpha of 5%.

The final randomised sample comprised 56 schools, 183 classes, and 5,074 children. There is data at baseline (September-October) and 12 months post-baseline for 4,477 children, and data at baseline and 24 months post-baseline for 4,147 children. There is data at all data collection points for 4,006 children.

Randomisation

Schools were randomised on a 1:1 ratio, stratified by two binary variables: school size and proportion of children eligible for free school meals. School size was set as large or small (n< 257 pupils) and the percentage of free school meals was set as high or low (< 39.8% pupils). The North Wales Organisation for Randomised Trials in Health (& Social Care) (NWORTH) prepared a computer generated random sequence. To ensure concealment, a web randomisation system was used to allocate schools. The allocation procedure was initiated by one of the research team at an event attended by representatives from the relevant schools.

Intervention condition

The pre-kindergarten version of PATHS was the intervention under trial. The programme developer recommended using this version because (a) it is designed for 4-5 year-olds, which overlaps with the age of children in the trial (those in Reception) and (b) the slightly older children (5-6 years at baseline) had not received PATHS lessons at a younger age and were therefore considered (by the programme developer) to need the more basic content. The programme is described extensively elsewhere (Greenberg & Kusché, 2002; Domitrovich et al., 2005). Briefly, it aims to develop skills in five main areas: self-awareness, managing feelings, motivation, empathy and social skills. The PATHS lesson materials and teacher guidance are contained in a series of manuals for teachers. PATHS lessons are developmentally sequenced to build on and integrate previous learning. They are delivered in classrooms by trained teachers and focus on: (i) techniques for self-control, (ii) emotional and interpersonal understanding, focusing specifically on

recognising and dealing with different feelings in oneself and others; (iii) steps for solving interpersonal problems; (iv) positive self-esteem; and (v) improved peer communication and relationships.

The lesson materials include teacher scripts, pictures, photographs, activity sheets, posters, home activities and parent letters and information. A supplementary booklet shows how these materials map onto UK curricular requirements and includes ideas for cross-curricular links (SRU, 2009). It also highlights how materials that were originally developed in the US need to be adjusted for teaching in schools in England. Suggested amendments, based on a thorough review of the materials by a local teacher, were surface level – for example, amending language, images and stories – and made in close collaboration with the programme developer so as to preserve the underlying logic. Lessons typically take up an hour per week and can be divided into two or more sessions to fit into the teaching programme of the class. There are 44 lessons in the Reception/Year 1 PATHS curriculum and 47 lessons in the Year 2 curriculum. PATHS is designed as a rolling programme, with lessons in Year 2 building upon the foundation provided in Year 1, and extending until the last year of primary/elementary school. However, it is also possible to deliver PATHS as a stand-alone programme with an individual year group if so desired. Additional whole school activities and strategies are suggested, as generalisation of the key messages throughout the school day is strongly encouraged. All school staff should be aware of and model the key PATHS ideas and routines. Activities are also included for parents to do with children at home.

Teachers in PATHS schools received one day of training by accredited PATHS trainers from the US prior to implementation (as recommended by the programme developer). In addition, they were supported by one of three coach consultants employed by the Council (all themselves trained teachers). Their role was to: assist teachers to deliver the programme with fidelity by co-ordinating training events (led by the programme developers); model PATHS lessons in-school where necessary; provide feedback and support to teachers, particularly on generalising PATHS techniques to other areas of teaching; and to be the school liaison on any matters related to the PATHS programme. The coach consultants were trained fully by the accredited PATHS trainers and received telephone, email and in-person support from these trainers throughout the implementation. Each coach was responsible for supporting about 10 schools on average, and sought to visit each school twice a term. This was deemed adequate but also realistic in the sense that a more intensive model would not be feasible in scale-up (see Bateman et al. 2012 for more about implementing PATHS in the city).

Intervention fidelity

Four dimensions of implementation fidelity were measured in this study. The first two, exposure/dose and adherence (to content), were measured via weekly logs, completed by teachers, that captured the delivery dates of each of the 44 or 47 PATHS lessons and whether they were ‘taught as written’, ‘taught with minor deviations from the curriculum’, or ‘taught with major deviations from the curriculum’. The third and fourth dimensions, quality of

programme delivery and pupil engagement respectively, were measured as part of a global rating of fidelity, using a semi-structured questionnaire completed by PATHS coaches in consultation with the teachers they were responsible for supporting. This measure also captured information on adherence to programme content and dose.

A low completion rate (27%) of the fidelity review questionnaire in the first year of implementation yielded insufficient data to report reliably on fidelity. Feedback from participating teachers at PATHS training events early in the second year of implementation suggested that they did not have enough time to complete the fidelity paperwork. The self-report lesson logs were shortened following this feedback and a higher response rate was obtained in the second year, with coach/teacher pairs completing 92% of the expected fidelity review questionnaires and teachers returning 78% of expected lesson logs. Fidelity results reported in this paper use this data from the second year of implementation.

Control condition

Twenty-seven schools were allocated to the waiting list control condition, continuing with services as usual: they were thus able to continue to use the national SEAL curriculum if they were already implementing it. Intervention schools were not required to stop SEAL but it is likely that most likely replaced SEAL with PATHS owing to curriculum space, at least for the classroom element. At the end of the two-year study (i.e. the 2011/2012 academic year) control schools were offered the opportunity to implement PATHS. As with

intervention schools, the city council covered the initial costs for training and materials if control schools took up this offer. Control schools were also offered an incentive of £1,000 each for taking part in all three data collection points for the study. Only two control schools withdrew from the study after baseline (see Appendix A). No other costs or expenses were paid to control schools or teachers.

Measures

Teachers completed the measures on all of the children in their class. This meant that in intervention schools the programme deliverer was also the main source of information about children's outcomes. Because children moved class grades over the two-year period of the study (from Reception to Year One, and from Year One to Year Two), the source of data on the children was not the same for all data collection points. For the most part, baseline and 12 months teachers were the same person but the 24 months teacher was more likely to be a different teacher than baseline. The study covered the cost of one half day of supply cover for classroom teachers at each data collection point in both intervention and control schools to ensure there was designated time to complete the research measures. In data collection points 12 and 24 months post-baseline, we entered all teachers who had contributed data into a lottery draw for 10 prizes of £50 book vouchers.

The primary outcome measure was the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997). Teachers rate their pupils' difficulties on 25 items. Scores are summed for four sub-scales where higher scores

indicate greater problems, namely conduct problems, emotional difficulties, hyperactivity, and peer relationships, and as well as one sub-scale where higher scores indicate better pro-social behaviour. In addition, a total difficulties score is calculated by summing the first four sub-scales. Cut-offs can be applied to provide an indication of likely clinical disorder (Goodman et al., 2000a, 2000b). A five-item impact scale indicates the extent of the burden that the child's problem behaviour has on peer relationships, classroom learning or the class as a whole. The SDQ displays good internal consistency (r = 0.73); re-test stability after four to six months (r = 0.62) and good discriminant validity demonstrated by high problems scores being associated with increased psychiatric risk (Goodman, 2001; Goodman and Scott, 1999).

A secondary measure of children's behaviour and social and emotional development was the *PATHS Teacher Rating Scale (PTRS)*, a series of standardised sub-scales as follows: (1) emotion regulation; (2) pro-social behaviour; (3) social competence; (4) aggressive behaviour; (5) internalising/withdrawn; (6) relational aggression; (7) peer relations; (8) inattention-hyperactivity; (9) impulsivity-hyperactivity; (10) learning behaviours; and (11) academic performance. Subscales 1-7 were rated on a six-point likert scale from "almost never" to "almost always", subscales 8-9 were rated on a four-point likert scale from "not at all" to "very much", and subscales 10-11 were rated on a three-point likert scale from "most often applies" to "does not apply". Four of the sub-scales are positively scored (i.e. higher score indicates better functioning), namely emotion regulation, pro-social behaviour, social competence, and learning behaviours. The remaining

sub-scales are negatively scored, where higher scores equate to greater problems. Teachers were also asked to indicate whether the child was “near the very bottom of the class; in the bottom half of the class; in the middle of the class; in the top half of the class; or near the very top of the class” for reading and literacy, maths and overall academic functioning (this formed the academic performance subscale).

Further, observations of child and teacher behaviours were undertaken by trained independent observers (two Psychology Researchers with PhDs) in classrooms in a third of the intervention and control schools (n =19, 10 intervention, 9 control) using the Teacher-Pupil Observation Tool (T-POT) (Martin et al., 2009). This 27-item measure covers the following nine domains: teacher positive behaviour (e.g. use of praise and encouragement); teacher negative behaviour (e.g. criticism, negatively phrased commands); teacher praise (specific and non-specific praise); class compliance / non-compliance (responses to teacher commands and questions); class negative behaviour to the teacher (e.g. verbal or physical aggression); class prosocial behaviour; class off-task behaviour; and ‘total negatives’ (sum of all negative behaviours by teachers and children). Observations were conducted in December 2009 (baseline) and June 2010 (i.e. 6 months post-baseline). Classrooms were observed for 20 minutes each. The observers were blind to condition.

In addition to the core outcome measures, teachers from PATHS and control schools were asked to provide demographic and professional information (e.g. behavioural management strategies) at baseline. The *Teacher*

Background questionnaire measures teachers' reports of their own well-being, perceptions of general school climate, orientation of the school to innovation, collective responsibility, and working with parents. These variables were used in the final model as classroom-level covariates or as school-level covariates where it was possible to combine the responses from teachers within the same school (e.g. school climate). In subsequent data collection points, a brief version of the questionnaire also contained items regarding the PATHS implementation support they had received from PATHS coaches.

Analysis

Statistical analyses were undertaken and reported in accordance with the CONSORT guidelines for cluster RCTs (Campbell et al., 2004). Analysis was based on the intention to treat (ITT) principle, according to the initial allocation of schools to ether intervention or control groups, and results are reported at the level of individual children. Intervention effects were estimated with mixed linear regression models, fitting both school and classroom as a random effect, with adjustment for baseline outcome, age, gender, ethnicity, qualifying for free school meals, special education needs as fixed factors.

Groups are compared at two points in time: baseline vs. 12-months post-baseline and baseline vs. 24-months post-baseline. As school classroom was the smallest cluster in the sampling design, an intra-class correlation (ICC) was calculated to compare the variation between school class and the total variance. Secondary analyses were undertaken to assess the impact of missing data and examine potential moderators of intervention effects. Eight

schools dropped out shortly after randomisation and collected no baseline children level data. For the remaining schools, multiple imputation was used to assess the impact of missing children level data at 12 and 24-months on the SDQ and PTRS (White et al., 2011).

The between-group difference inferences of imputed models were contrasted to complete case ITT models. In order to assess the level of potential attrition/loss to follow-up, the characteristics of the children were compared according to those with and without missing data. Interaction tests were performed for a number of predefined children level covariates on the primary outcome of SDQ total score: gender, age, ethnicity, qualification for free school meals, special education needs and baseline SDQ score, school level covariates (school size, provision of SEAL) and cohort. These interaction tests were undertaken using imputed data sets in order to maximise statistical power (Little & Rubin, 2002).

All statistical tests were two-sided, and deemed to be statistically significant if $p \leq 0.05$. No adjustments for multiplicity were made because the outcome variables were inter-related. Simple adjustments for the number of comparisons were overly conservative (Brown & Russell, 1997). The focus of the data presentation is the estimation of the mean between-group difference and 95 percent confidence interval. We also report the ICC. All analyses were performed with Stata version 11.0 (Statacorp, College Station, TX, USA).

In the case of the T-POT, the data was analysed using ANCOVA, focusing on the difference between baseline (which as indicated earlier was 1-2 months after the intervention commenced) and six months post-baseline. Two-thirds of the 27 variables were not normally distributed, so in order to increase the power and ensure that analyses were performed on normally distributed data, variables were combined into nine composite categories (corresponding to the domains referred to above). Baseline data for one intervention school was carried forward to the 6 months data point because it was not possible to observe the classroom at 6 months. Owing to the T-POT only being applied to a sub-sample, the results need to be taken with caution and are not definitive.

A cost consequence analysis was undertaken alongside the RCT, which involved measuring programme costs from a multi-agency public sector perspective (Edwards et al., 2008) alongside the measurement of child behaviour outcomes. Costs were reported in UK £ for the year 2009-10. Discounting was not required, as study completion was within one year.[4] Micro-costing methodology was applied when costing the intervention (Curtis & Netton, 2004; Drummond et al., 2005; Edwards et al., 2007).[5]

[4] Discounting helps to 'make fair' comparisons of programmes whose costs and outcomes occur at different times. In order to compare, and add up, costs that accrue at different times it is necessary to calculate their present value, which expresses them as an equivalent amount of today's pounds. Discounting converts the £ value of costs in different time periods to present value. This is necessary because a £ in the future is worth less than a £ now.

[5] Micro-costing is bottom-up approach used to estimate the cost of setting up and delivering an intervention. It involves collecting detailed information about the resources required to deliver an intervention, and subsequently assigning economic unit costs to each component of resource use. The alternative approach would be gross-costing, a top-down approach where the total cost invoiced is divided by the total resource use to obtain an average cost of resource use. The micro-costing approach is accepted as being more accurate than gross-costing, and is widely used in costing studies (e.g. Tarricone, 2006; Charles et al., 2013; Xu et al., 2014).

The costs of the PATHS programme, including training, administration and delivery, were obtained from the developers and the city council's 'Brighter Futures' team. The time and resources recommended by developers to prepare and deliver the programme in schools was verified by direct contact with four champion teachers (teachers with responsibility in their schools for ensuring that PATHS was delivered well). Teachers' costs were based on national average salaries for a mid point M4-6 qualified teacher. A school year of 38 weeks for delivery of PATHS was assumed, taking sickness, training (CPD – Continuing Professional Development) and holidays into account salary in calculations, inclusive of employers' on-costs (25%). The training teachers received in PATHS will accrue a future benefit to reflect this. The costs were annuitised over 5 years at 3.5% (Netten & Knight, 1999).

The PATHS programme costs were separated into recurrent and non-recurrent costs. The various elements were summed and divided by the number of schools delivering PATHS (n=29) to give a cost per school and the cost per child, with total cost divided by the number of children receiving the intervention (N = 2640). The recurrent costs are also presented separately by school and by pupil. These costs were to be used to establish cost per increment drop on the teacher-completed SDQ (Goodman, 1997).

Results

The CONSORT diagram in Appendix A depicts the flow of recruitment and retention to the trial, at the school, classroom and individual child level, as recommended by best practice for RCTs (Schulz et al., 2010). Sixty-four

schools were randomised in July 2009. Eight schools dropped out immediately following randomisation (two PATHS and six control schools) and therefore did not collect baseline data on their children. Nevertheless, the remaining 56 schools showed a good balance between PATHS and control groups in terms of the stratification variables (Table 1).

Table 1 about here

Baseline characteristics

There were no statistically significant differences between the programme and control group at baseline on key demographic factors or the primary outcome measures (Table 2).

Table 2 about here

Outcomes

The ITT analysis for the primary child outcome measure (SDQ) was undertaken on all children for whom there was baseline data. Models were fitted for those children for whom there was complete data for the two time points compared. Two different imputation methods (Last Observation Carried Forward (LOCF) and multiple imputation) were used to handle missing data. The different methods did not result in substantive differences in the findings, although imputation models were more likely to result in non-significant effects.

Table 3 presents the results from the Hierarchical Linear Modelling (HLM) analysis for the SDQ, where the difference in outcomes between the PATHS group and control group is examined, accounting for class-level and school-level clustering in outcomes and available covariates at all levels. The between-group mean difference, 95% confidence intervals and ICC values, are presented for the complete case analysis. The imputed model using multiple imputation methods is presented in Table 3. Table 4 presents the same data for the secondary outcome measure, the PTRS.

Tables 3 and 4 about here

Regarding the primary outcome measure (SDQ), there were no statistically significant differences at 12 months post-baseline for the total difficulties and impact scores and each of the five subscales. For the PTRS, at 12 months there were statistically significant differences favouring the programme group for six out of 11 sub-scales: social competence, aggressive behaviour, inattention-hyperactivity, impulsivity-hyperactivity, peer relations and learning behaviours.

Independent observations of teacher and classroom behaviour using the T-POT instrument (Table 5) revealed statistically significant ($p<0.05$) differences favouring the intervention group for only three of the nine composite categories, all with effect sizes towards the lower end of the range identified by Durlak et al. (2011): teacher total positive behaviours (.304); class behaviour negative to teacher (.307); and class off-task behaviour (.227).

Non-significant changes in the right direction were found for a further three categories: teacher negative behaviours; class pro-social behaviour; and class total negative behaviours.

Table 5 about here

However, as Tables 3 and 4 indicate, at 24 months post-baseline, gains on the PTRS had been lost – there were no statistically significant differences at 24 months for any of the PTRS subscales, either for complete case analysis or the imputed model – and there were still no statistically significant differences on the SDQ (using complete case analysis, the control group performed better for all but two of the five sub-scales – peer relations and prosocial behaviour). Using the imputed model there was a statistically significant difference for conduct at 24 months post-baseline, favouring the control group (Table 3).

There were some sub-group differences (Table 6). There was a significant impact at 24 months on pupils who tested as having emotional difficulties at baseline on the SDQ. These pupils in PATHS schools were significantly more likely to show improvement on the overall score of total difficulties on the SDQ compared with similar pupils in control schools. White pupils benefited more than other ethnic groups, though not significantly so. Age and poverty did not emerge as significant moderators of the results.

Table 6 about here

Fidelity of implementation

PATHS lessons were designed to be delivered in sequence, exposing children to two or more lessons per week. In terms of *dose*, by the end of year two, according to teacher-completed records (response rate of 78%), PATHS teachers had completed an average of 26 lessons (55% of the 47 possible lessons). Coaches reported a higher rate of lessons delivered than records suggest. There was significant variation across the PATHS schools, with some teachers only delivering up to lesson eight while others had delivered all 47 lessons.

In terms of *adherence*, levels of self-reported teacher adherence to PATHS lesson content were high. On average, teachers reported that lessons were delivered 'as specified' in the programme materials. Almost half of the children in the intervention schools participated in lessons that were delivered 'as intended/written' (44% of all children in PATHS schools), with a further 54% participating in lessons that were delivered by teachers with only *minor* deviations from the model (e.g. using a different book suggestion or with some reference to personal experience). Very small numbers of children attended a lesson with major deviations (< 2%).

By contrast, there was variation in the global rating of PATHS delivery, which included aspects of quality such as the extent to which teachers were familiar with curriculum content and could generalise outside of the PATHS lessons. Coaches' overall ratings of lesson delivery (expressed as a percentage of the

total possible score on the fidelity reviews) ranged from 21% to 100%, with a mean fidelity score for the sample of teachers of 79% (SD = 21%). Drawing an arbitrary threshold of 80% implementation fidelity[6], 47 out of a possible 94 PATHS teachers (50%) could be said to have delivered the programme with 'high fidelity'.

Outcome by fidelity

Given that the intervention showed no effect, it is appropriate to examine whether a possible explanation is fidelity of delivery. As described in the previous section, implementation fidelity according to the global rating was high on average in the PATHS schools (mean = 79%). Schools were divided into two categories according to whether they were above or below this average figure (calculated as an average of teachers in the school). A comparison of 'high' (80% or greater fidelity as rated by coach/teacher pairs) and 'low' (<79%) fidelity schools indicated that although the social and behavioural outcomes were better for children in schools that implemented PATHS well, 'high' fidelity schools still do not significantly outperform control schools in a GLM model fitted using only high fidelity schools as the PATHS condition predicting treatment effects on child outcomes.

Costs

Given the lack of impact on child outcomes, cost-effectiveness of PATHS was not calculated. The micro-costing analysis is summarised in Table 7. Non-recurring costs associated with *training* 101 teachers to deliver PATHS

[6] There is no evidence-based threshold for fidelity advised by the programme developer. The 80% threshold used here is purely indicative and other intervention studies have set lower thresholds for acceptable and high fidelity.

programme were £140,501 (£1,391 per teacher). The recurring costs associated with the *management* and *delivery* of the PATHS programme in each of the 29 schools was £11,593 (£127 per child, based on delivery to 2,640 children). (This is a sum of the amount for management, which is £40,950 across 29 schools, or £1,412 per school, and the amount for delivery, which is £295,249 across 29 schools, or £10,181 per school.) When the non-recurring costs of initial training are annuitised over five years, based on the UK teacher salary band (4-6), the costs per school were £12,666 and the costs per child £139.

Table 7 about here

Discussion

Overall, the evaluation showed little or no benefit for child social-emotional outcomes (including behaviour) following the introduction of the PATHS intervention compared to control schools. At both 12 months and 24 months post-baseline there were no statistically significant effects favouring the PATHS group on the primary outcome measure (SDQ) and the independent observation of classroom behaviour at six months post-baseline did not indicate sufficient significant change to be considered clinically meaningful. The average cost of the PATHS intervention was £12,666 per school or £139 per child (assuming an average of 91 children receiving the intervention per school).

There were some important sub-group differences. Echoing studies of the

PATHS programme elsewhere (Kam et al., 2004), children identified at the outset as having emotional difficulties (based on reaching the clinical cut-off on the SDQ emotions subscale) did significantly better in PATHS schools than in controls. This speaks to the value of targeted programmes for this age group and subject area, an argument also advanced in relation to school-based programmes to prevent depression (Spence & Shortt, 2007), but this may be harder to deliver, at least in a school setting. The city where the study took place has a large proportion of children and young people are from minority ethnic groups (approximately 37%), although pupils of Pakistani origin benefited no more than 'white' pupils from PATHS. Poverty and age also did not emerge as significant moderators of results.

Previous European PATHS evaluations (Malti et al., 2011; Ross et al., 2012) did not include an economic evaluation with which to compare our cost findings associated with introducing the programme into schools. Despite the unexpected findings, this paper has provided much needed evidence as to the outcome and costs associated with delivering PATHS in such a large and ethnically diverse city (population figures from 2011: White British 53.1%, Pakistani 13.5%, Indian 6%, White other 4.8%, Caribbean 4.4%, Mixed 4.4%, Bangladeshi 3%, African 2.8%, Chinese 1.2%, Other ethnicity 6.7%).

Given the previous weight of evidence behind PATHS in particular, and SEL programmes in general, we propose five hypotheses for the less than positive finding on child outcomes in this trial.

First, it is possible that PATHS is less effective in those European countries in which there is regular SEL provision than in countries such as the US where such provision is arguably weaker. The results from the present study are similar to those found by Malti et al. (2011) at a comparable post-test in a large trial of PATHS in Zurich, Switzerland, although their longer-term follow-up did indicate that some benefits may only become apparent later. Ross et al. (2012) found small significant benefits in Northern Ireland for a PATHS experiment in its first year but still well short of those reported in the US. In both studies the authors suggested the strength of existing provision as a reason for the disappointing impact, and it has been cited as an explanation for the lack of effect of other US-origin programmes in some European countries, albeit in different areas of children's services (e.g. Sundell et al., 2008). However, the plausibility of this hypothesis is questionable. To start with, the amount, nature and quality of Personal, Social and Health Education (PSHE) education (which covers SEL) in schools in England are variable (Ofsted, 2013; Formby et al., 2011). In the present study it is known that a significant proportion of control schools were implementing SEAL, which, as indicated, is a government-sponsored programme with overlapping aims and themes with PATHS. However, the evidence for the effectiveness of SEAL is equivocal at best. Specifically, Humphrey et al.'s (2010) quasi-experimental national evaluation of SEAL in secondary schools found no impact on child outcomes, while the evaluation of primary school SEAL found 'a complicated picture of impact' with some positive effects countered by negative, possibly iatrogenic effects (Humphrey et al. 2008: 6). Further, there is no other evidence that US-origin SEL programmes fail in Europe because, to our

knowledge, no other such programme (i.e. besides PATHS) has been evaluated experimentally in Europe.

A second possibility is that with relatively small effect sizes expected (i.e. at the lower end of the range identified by Durlak et al. (2011)), even small losses in fidelity of implementation can eradicate impact on child outcomes. However, as the results indicate, fidelity for PATHS in this study was reasonable when measured in terms of implementation quality, particularly in the second year of implementation. Moreover, there was no measurable relationship between fidelity and treatment effects: 'high fidelity' (as measured using an arbitrary 80% fidelity threshold) PATHS schools outperformed lower fidelity ones but not controls.

A third hypothesis concerns the timing of the measurement of child outcomes. Both the Zurich study (Malti et al., 2011) and the US study (Riggs et al., 2006) found no significant effects of PATHS on children's conduct problems at immediate post-test, the comparable measurement point for this evaluation. They did, however, find significant effects at one and two years post-intervention, indicating that it may be a matter of time before universal prevention programmes like PATHS demonstrate their benefits. This is because rather than necessarily immediately enhancing children's social competence, the skills conferred by the intervention may only become needed or demonstrable at a later time. Unfortunately, funding for this evaluation did not stretch to a long-term follow-up with participants. Moreover, the fact that effects on six of the 11 PTRS sub-scales at 12 months post-baseline had

disappeared by 24 months post-baseline suggests that the possibility of 'sleeper effects' in this instance is highly unlikely.

Fourth, it is important to consider measurement error. Re-analysis of the data revealed no cohort or teacher reporting effects. Data were examined for acquiescent responding and significant changes in the variance (of scores) over the three data collection points. There were no indications from these analyses that measurement error was behind the lack of significant findings. However, as indicated earlier, the teacher-deliverers were also the sole source of data on the children's outcomes.

Finally, it is possible that PATHS teachers, trained in social-emotional development, became more sensitive over time to children's difficulties and, as a result, were more likely to report problems compared to control teachers. In addition, this led to high intra-class correlations because it accounted for both the true similarity of children within a class as well as some characteristic or personal quality of the teacher (e.g. a tendency to view most children in their class as either good or bad). Ideally, it would have been possible to supplement the teachers' reports with those of pupils or parents, but this could not be done with the resources available. It is also worth noting that studies that have gathered data from more than one source often report significant variation between sources, making conclusions difficult (e.g. Malti et al., 2011; Humphrey et al., 2010).

This study has considerable strengths, including that it was fully powered and

independent of the programme developers. However, several limitations should be noted. First, SEAL was possibly still being implemented in some schools but this was not monitored fully. National SEAL monitoring data returns were only available for a handful of the participating schools. Second, completion rates of fidelity forms in the first year of the study in particular were low. Third, the main outcome measure was completed by teachers who also delivered the programme. Although expensive, independent observation may be needed for the whole sample to provide a more robust indication of effect. Fourth, the first T-POT data collection point was approximately 1-2 months after the programme started (i.e. not a true pre-intervention baseline measure). Finally, of the 18 teachers observed using the T-POT at six months post-baseline, only 10 were the same as those observed at baseline due to normal staff turnover in the schools. In other words, almost half the teachers were different at the later data collection point, meaning that it is not possible confidently to attribute changes on the T-POT to the effect of PATHS; they may simply be the effect of the classroom interacting differently with the teacher at the second observation.

Conclusions

This is first fully powered independent RCT of PATHS in the UK, and it shows little or no effect of the intervention on child behaviour. Although this study has failed to replicate findings from previous PATHS evaluations or to add to the existing evidence in support of SEL programmes in general, it has provided important data on the realistic fidelity of implementation for an evidence-based intervention in UK schools as well as the cost of setting up and delivering

such an intervention. While it was not possible to conduct a cost-effectiveness analysis owing to the lack of impact, a recent analysis based on present value life cycle benefits and costs calculated a benefit-cost ratio for PATHS of $15.66; importantly, however, given the results of the present study, the benefits were all derived from increased earnings as a results of better academic test scores, not from reduced health or crime costs owing to improved behaviour (the meta-analysis found no statistically significant effects on behaviour) (WSIPP, 2015). In terms of implications for policy and practice, the study shows that it cannot be assumed that an evidence-based programme will work in all contexts and that testing in local settings is important.

OTHER INFORMATION:

Trial registration site and number

www.controlled-trials.com: ISRCTN 32534848

Funding: This study was funded by Birmingham City Council (BCC) as part of its 'Brighter Futures' strategy for children's services.

Acknowledgements

The authors would like to acknowledge the hard work and dedication of Birmingham City Council's (BCC) 'Brighter Futures' team, including the PATHS coaches who collected fidelity data for the study. In addition, we thank the teachers and schools that took part in the trial, contributing valuable data on children's well-being. Professor Daniel F. Perkins was involved in the design of the study. Drs Pam Martin and Catrin Eames undertook the independent observations of classroom behaviour. Pat Linck undertook initial analysis of the cost data. The North Wales Trials Unit (NWORTH) provided the randomisation service for this study.

Interests declared: None

Ethical clearance: The Warren House Group Ethics Committee.

REFERENCES:

American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)*, Amercian Psychiatric Association.

Bateman, A., Boyes, S., Hines, J., Lacey, T., Kerrigan, B., Hopkins, C., Lehtonen, M. & Axford, N. (2012). Implementing the PATHS program in Birmingham, England. In Brown, P. M., Corrigan, M. W. & Higgins-D'Alessandro, A. (Eds.) *Handbook of Prosocial Education*. Blueridge Summit PA: Rowman and Littlefield.

Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., Blair, C., ... Gill, S. (2008). Promoting Academic and Social-Emotional School Readiness: The Head Start REDI Program. *Child Development*, 79(6), 1802-1817.

Blueprints (2014) *Promoting Alternative Thinking Strategies (PATHS)* http://www.blueprintsprograms.com/factSheet.php?pid=b6692ea5df920cad691c20319a6fffd7a4a766b8 [Accessed 3rd January 2014]

Blueprints (2012). *Blueprints for Violence Prevention* website. http://www.colorado.edu/cspv/blueprints/modelprograms/PATHS.html. Accessed 26th September 2012.

Brown, B. and Russell, K. (2007). Methods correcting for multiple testing:

operating characteristics. *Statistics in Medicine,*16, 2511-28.

Bywater, T. & **Sharples, J.** (2012). Effective evidence-based interventions for emotional well-being: lessons for policy and practice. *Research Papers in Education* 27(4), 389–408.

Campbell, M.K., Elbourne, D.R., Altman, D.G. for the CONSORT Group. (2004). The CONSORT statement: extension to cluster randomised trials. *British Medical Journal*, 328, 702-708.

CASEL (Collaborative for Academic, Social and Emotional Learning) (2014) http://www.casel.org/social-and-emotional-learning [Accessed 2nd January 2014]

Charles, J. M., Edwards, R. T., Bywater, T., & Hutchings, J. (2013). Micro-costing in public health economics: Steps towards a standardized framework, using the Incredible Years Toddler Parenting Program as a worked example. *Prevention Science*, 14(4), 377-389.

CPPRG (Conduct Problems Prevention Research Group) (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology*, 67(5), 648-57.

CPPRG (2010). The effects of a multiyear universal social-emotional learning program: the role of student and school characteristics. *Journal of Consulting*

and Clinical Psychology, 78(2), 156-168.

Curtis, C. & Norgate, R. (2007). An evaluation of the PATHS curriculum at Key Stage 1. *Educational Psychology in Practice* 23, 33-44.

Conduct Problems Prevention Research Group (CPPRG) (1999). Initial impact of the Fast Track prevention trial for conduct problems: II. Classroom effects. *Journal of Consulting and Clinical Psychology,* 67(5), 648-57.

Curtis, L. & Netten, A. (2004). *Unit Costs of Health and Social Care 2004*. Canterbury: PSRRU.

Denham, S. A. & Brown, C. (2010). “Plays nice with others”: social-emotional learning and academic success. *Early Education and Development* 21(5), 652-680.

DfES (Department for Education and Skills) (2005) *Primary Social and Emotional Aspects of Learning (SEAL): Guidance for Schools*. Nottingham: DfES Publications.

DfES (Department for Education and Skills) (2007) *Social and Emotional Aspects of Learning (SEAL): Guidance for Secondary Schools*. Nottingham: DfES Publications.

Domitrovich, C., Cortes, R., & Greenberg, M. (2007). Improving young

children's social and emotional competence: a randomized trial of the preschool "PATHS" Curriculum. *Journal of Primary Prevention*, 28(2), 67-91.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development,* 82(1), 405-432.

Domitrovich, C. E., Greenberg, M. T., Cortes, R., & Kusché, C. A. (2005). *The Preschool PATHS Curriculum.* Deerfield, MA: Channing-Bete Publishers.

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'Brien, B.J., & Stoddart, G.L. (2005). *Methods for the Economic Evaluation of Health Care Programmes (3rd Ed.)*. Oxford: Oxford University Press.

Ecclestone, K. & Hayes, D. (2008) *The Dangerous Rise of Therapeutic Education*. London: Routledge.

Edwards, R.T., Ó Céilleachair, A., Bywater, T., Hughes, D.A. and Hutchings, J. (2007). Parenting programme for parents of children at risk of developing conduct disorder: cost effectiveness analysis. *British Medical Journal* 10, 334-682.

Edwards, R.T., Hounsome, B., Linck, P., & Russell, I.T. (2008). Economic evaluation alongside pragmatic randomised trials: developing a Standard Operating Procedure for clinical trials units*. Trials,*14, 64-65.

Formby, E., Coldwell, M., Stiell, B., Demack, S., Stevens, A., Shipton, L., Wolstenholme, C. & Willis, B. (2011) *Personal, Social, Health and Economic (PSHE) Education: A Mapping Study of the Prevalent Models of delivery and their Effectiveness*. Research Brief DFE-RB080. London: Department for Education.

Furedi, F. (2009) *Wasted: Why Education Isn't Educating*. London: Continuum.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.

Goodman R (2001) Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ). *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 1337-1345

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000a). Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *British Journal of Psychiatry*, 177, 534-539.

Goodman R., Renfrew D., & Mullick M. (2000b). Predicting type of psychiatric disorder from Strengths and Difficulties Questionnaire (SDQ) scores in child mental health clinics in London and Dhaka. *European Child and Adolescent Psychiatry,* 9, 129-134.

Goodman, R. & Scott, S. (2009). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Child Psychology* 27 (1), 17-24.

Greenberg, M. T., & Kusché, C. A. (1998). Preventive intervention for school-age deaf children: the PATHS curriculum. *Journal of Deaf Studies and Deaf Education*, 3(1), 49-63.

Greenberg, M. T., & Kusché, C. A. (2002). *Promoting Alternative Thinking Strategies (PATHS): Blueprints for Violence Prevention (Book 10)*. Boulder, CO: Center for the Study and Prevention of Violence, University of Colorado.

Hobbs, T., Axford, N. & Jodrell, D. (2011). Getting the measure of child health and development outcomes (2): The picture for a local authority in England, *Child Indicators Research* 4(1), 81-100.

Hoffman, D. M. (2009). Reflecting on social emotional learning: a critical perspective on trends in the United States. *Review of Educational Research* 79 (2), 533-556.

Humphrey, N. (2013). *Social and Emotional Learning: A Critical Appraisal*. London: Sage.

Humphrey, N., Kalambouka, A., Bolton, J., Lendrum, A., Wigelsworth, M., Lennie, C., & Farrell, P. (2008). *Primary Social and Emotional Aspects of Learning (SEAL): Evaluation of Small Group Work*. London: Department for Children, Schools and Families.

Humphrey, N., Lendrum, A., & Wigelsworth, M. (2010). *Social and Emotional Aspects of Learning (SEAL) Programme in Secondary Schools: National Evaluation*. London: Department for Education.

Kam, C.-M., Greenberg, M. T., & Kusché, C. A. (2004). Sustained effects of the PATHS curriculum on the social and psychological adjustment of children in special education. *Journal of Emotional and Behavioral Disorders,* 12, 66-78.

Little, M., Berry, V., Morpeth, L., Blower, S., Axford, N., Taylor, R., Bywater, T., Lehtonen, M. & Tobin, K. (2012). The impact of three evidence-based programmes delivered in public systems in Birmingham, UK, *International Journal of Conflict and Violence* 6 (2), 260-272.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical Analysis with Missing Data* (2nd ed.). Hoboken NJ: John Wiley & Sons.

Malti, T., Ribeaud, D., & Eisner, M. P. (2011). The effectiveness of two universal preventive interventions in reducing children's externalizing behavior: A cluster randomized controlled trial. *Journal of Clinical Child and Adolescent Psychology,* 40(5), 677-692.

Martin, P. A., Daley, D., Hutchings, J., Jones, K., Eames, C., & Whitaker, C. J. (2010). The Teacher-Pupil Observation Tool (T-POT). *School Psychology International,* 31(3), 229-249.

Netten, A. & Knight, J. (1999). Annuitizing the human capital investment costs of health service professionals. *Health Economics*, 8, 245-255.

Ofsted (2011). *Not Yet Good Enough: Personal, Social, Health and Economic Education in Schools*. London: Ofsted.

Riggs, N. R., Greenberg, M. T., Kusché, C. A., & Pentz, M. A. (2006). The mediational role of neurocognition in the behavioral outcomes of a social-emotional prevention program in elementary school students: Effects of the PATHS Curriculum. *Prevention Science*, 7, 91-102.

Ross, S. M., Cheung, A., Slavin, R., Sheard, M. K., & Elliott, L. (2011). Promoting primary pupils' social-emotional learning and pro-social behaviour: Longitudinal evaluation of the Together 4 All Programme in Northern Ireland. *Effective Education,* 3, 61-81.

Schulz, K.F., Altman, D.G. & Moher, D., for the CONSORT Group. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, 23, 340:d40.

Seifer, R., Gouley, K., Miller, A. L., & Zakriski, A. (2004). Implementation of the PATHS curriculum in an urban Elementary School. *Early Education and Development*, 15(4), 471-486.

Spence, S. H., & Shortt, A. L. (2007). Research review: can we justify the widespread dissemination of universal, school-based interventions for the prevention of depression among children and adolescents? *Journal of Child Psychology and Psychiatry* 48 (6), 526-542.

The Social Research Unit (SRU) (2009). *A Service Design Manual for Implementing PATHS in Birmingham, UK*. Dartington: The Social Research Unit.

StataCorp. (2009). *Stata Statistical Software: Release 11*. College Station, TX: StataCorp LP.

Sundell, K., Hansson, K., Löfholm, C. A., Olsson, T., Gustle, L. H. & Kedesjö, C. (2008) The transportability of multisystemic therapy to Sweden: short-term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology* 22 (4), 550-560.

Tarricone, R. (2006). Cost-of-illness analysis: what room in health economics? *Health Policy*, 77(1), 51-63.

Weare, K. (2004). *Developing the Emotionally Literate School*. London: Sage.

Weare, K. (2010). Mental health and social and emotional learning: evidence, principles, tensions, balances. *Advances in School Mental Health Promotion* 3 (1), 5-17.

Weare, K. & Gray, G. (2003) *What Works in Promoting Children's Emotional and Social Competence and Wellbeing?* Nottingham: DfES Publications.

White, I. R., Horton, N. J., Carpenter, J., & Pocock, S. J. (2011). Strategy for intention to treat analysis in randomised trials with missing outcome data. *British Medical Journal,* 342:d40.

WSIPP (Washington State Institute for Public Policy) (2015). *Promoting Alternative Thinking Strategies (PATHS)*. http://www.wsipp.wa.gov/BenefitCost/Program/94.

Xu, X., Nardini, H. K. G., & Ruger, J. P. (2014). Micro-costing studies in the health and medical literature: protocol for a systematic review. *Systematic Reviews*, 3(1), 47.

Table 1

Characteristics of 56 Randomised Schools

Characteristics	PATHS schools (n=29)	Control schools (n=27)
School Size		
Large (i.e. >= 257 pupils)	16 (55%)	14 (52%)
Small (i.e. < 257 pupils)	13 (45%)	13 (48%)
% Free School Meals		
High (i.e. >= 39.8% of pupils)	15 (52%)	13 (48%)
Low (I.e. < 39.8% of pupils	14 (48%)	14 (52%)

Table 2

Baseline Demographics of the Children in Intervention and Control Schools

Demographic variables	Control N = 1,801	PATHS N = 2,203	Total N = 4,004
Percentage boys (count)	52% (937)	49% (1079)	0% (2002)
Percentage with SEN^ (count)	26.6% (479)	22.1% (487)	24.1% (965)
Mean age (SD)	5.08 (0.59)	5.06 (0.57)	5.07 (0.58)
Percentage non-white (count)	67.5% (1216)	68.6% (1511)	68.1% (2727)
Percentage > behaviour clinical cut-off * (count)	6.5% (117)	8.3% (183)	7.5% (300)
Percentage > emotions clinical cut-off * (count)	3.7% (67)	4.6% (101)	4.2% (168)
Percentage > total difficulties clinical cut-off * (count)	8.5% (153)	11.3% (249)	10% (400)

Note. * Cut-off = score of four or more on the SDQ behaviour, six or more on emotional difficulties, 16 or more on total difficulties, ^ SEN defined as either low (school action only), medium (school action plus), or high (statemented)

Table 3

Strengths and Difficulties Questionnaire[7] Results at Baseline and Follow Up, Between-Group Differences and Imputed Analyses[8]

SDQ Subscale	Baseline (pre-intervention) N Mean (SD)		12 months post-baseline N Mean (SD)		24 months post-baseline N mean (SD)		Complete case, Between-group difference, Mean (95% CI)* ICC+		Multiple Imputation, Between-group difference, Mean (95% CI)	
	PATHS	Control	PATHS	Control	PATHS	Control	12 months adjusted for baseline	24 months adjusted for baseline	12-months adjusted for baseline	24 months adjusted for baseline
Total score	N=2626 7.54 (6.06)	N=2380 6.79 (5.81)	N=257 2 6.15 (5.72)	N=2302 6.16 (5.74)	N=250 9 6.60 (6.24)	N=2168 6.39 (6.05)	N=4255 -0.42 (-1.11 to 0.28) ICC 0.25	N=3934 0.19 (-0.64 to 1.03) ICC 0.25	N=5542 -0.36 (-0.93 to 0.22)	N=5540 0.08 (-0.58 to 0.75)
Impact	N=2555 0.37 (0.94)	N=2353 0.30 (0.88)	N=255 0 0.26 (0.82)	N=2268 0.29 (0.88)	N=259 8 0.33 (0.92)	N=2276 0.32 (0.90)	N=4123 -0.03 (-0.11 to 0.05) ICC 0.09	N=4000 0.04 (-0.05 to 0.12) ICC 0.08	N=5523 -0.04 (-0,11 to 0.01)	N=4221 0.06 (-0.02 to 0.14)
Conduct	N=2628 0.99 (1.62)	N=2382 0.89 (1.51)	N=257 5 0.86	N=2305 0.93 (1.66)	N=251 9 1.14	N=2177 1.04 (1.80)	N=4265 -0.15 (-0.31 to 0.01) ICC 0.12	N=3953 0.16 (-0.04 to 0.35) ICC 0.12	N=5542 -0.12 (-0.27 to 0.02)	N=5227 0.19 (0.02 to 0.35)*

[7] 3-level (school, class, child) hierarchical model with random effects model and adjusting for baseline outcome, age, gender, ethnicity, qualifying for free school meals, special education needs and for school-level stratification variables i.e. size of school, proportion of free school meals
Negative score indicates follow up score in intervention group < control group +ICC for classroom

[8] 3-level (school, class, child) hierarchical model with random effects model and adjusting for baseline outcome, age, gender, ethnicity, qualifying for free school meals, special education needs and for school level stratification variables i.e. size of school, proportion of free school meals and for school level stratification variables i.e. size of school, proportion of free school meals. Negative score indicates follow-up score in intervention group < control group

			(1.56)		(1.92)					
Emotion	N=2628	N=2382	N=257	N=2305	N=251	N=2177	N=4265	N=3953	N=5542	N=4098
	1.25	1.06	5	1.09	8	1.20	-0.12 (-0.33 to 0.10)	0.06 (-0.18 to 0.30)	-0.10	0.10
	(1.95)	(1.95)	1.05	(1.74)	1.26	(1.83)	ICC 0.18	ICC 0.16	(-0.28 to 0.08)	(-0.14 to 0.35)
			(1.74)		(1.86)					
Pro-social	N=2627	N=2383	N=257	N=2305	N=252	N=2177	N=4265	N=3953	N=5542	N=5527
	6.60	6.86	6	7.50	0	7.56	0.18 (-0.16 to 0.52)	0.16 (-0.27 to 0.59)	0.19	0.14
	(2.70)	(2.64)	7.59	(2.53)	7.78	(2.49)	ICC 0.26	ICC 0.27	(-0.08 to 0.47)	(-0.16 to 0.45)
			(2.49)		(2.38)					
Hyper-activity	N=2628	N=2381	N=257	N=2305	N=251	N=2177	N=4263	N=3951	N=5542	N=5527
	3.59	3.31	5	2.88	8	2.89	-0.03 (-0.32 to 0.26)	0.09 (-0.24 to 0.42)	-0.08	0.13
	(3.08)	(2.91)	2.95	(2.89)	3.00	(2.86)	ICC 0.15	ICC 0.17	(-0.27 to 0.17)	(-0.14 to 0.39)
			(2.93)		(2.91)					
Peer relation-ships	N=2628	N=2382	N=230	N=2576	N=251	N=2177	N=4265	N=3953	N=5542	N=4099
	1.70	1.52	5	1.28	8	1.24	-0.03 (-0.26 to 0.20)	-0.03 (-0.28 to	-0.02	0.01
	(1.82)	(1.78)	1.25	(1.60)	1.21	(1.62)	ICC 0.140	0.22)	(-0.18 to 0.13)	(-0.21 to 0.36)
			(1.58)		(1.65)			ICC 0.23		

Note. * indicates $p < 0.05$

Table 4

PTRS Questionnaire[9] Results at Baseline and Follow Up, Between-Group Differences and Imputed Analyses[10]

PTRS subscale	Baseline (pre-intervention) N Mean (SD)		12 months post-baseline N Mean (SD)		24 months post-baseline N mean (SD)		Complete case, Between-group difference, Mean (95% CI)*		Multiple Imputation, Between-group difference, Mean (95% CI)*	
	PATHS	Control	PATHS	Control	PATHS	Control	12 months adjusted for baseline	24 months adjusted for baseline	12 months adjusted for baseline	24 months adjusted for baseline
Emotion regulation	N = 2584 4.22 (1.00)	N = 2294 4.28 (0.95)	N = 2585 4.42 (1.02)	N = 2319 4.37 (1.00)	N = 2627 4.27 (0.99)	N = 2295 4.45 (1.02)	N=4203 0.11 (-0.04 to 0.27) ICC 0.31	N=4019 -0.18 (-0.35 to 0.00) ICC 0.35	N = 5539 0.13 (-0.02 to 0.28)	N = 5533 -0.18 (-0.35 to -0.01)
Pro-social behaviour	N = 2584 3.91 (1.11)	N = 2294 3.98 (1.08)	N = 2585 4.35 (1.11)	N = 2319 4.23 (1.11)	N = 2627 4.32 (1.07)	N = 2295 4.36 (1.12)	N=4203 0.16 (-0.01 to 0.32) ICC 0.37	N=4019 -0.06 (-0.25 to 0.13) ICC 0.37	N = 4876 0.20 (0.04 to 0.35)	N = 4876 -0.06 (-0.24 to 0.14)
Social competence	N = 2584 4.05 (0.99)	N = 2294 4.12 (0.96)	N = 2585 4.38 (1.01)	N = 2319 4.29 (1.01)	N = 2627 4.30 (0.98)	N = 2295 4.40 (1.02)	N=4203 0.14 (0.01 to 0.29)* ICC 0.38	N=4019 -0.11 (-0.29 to 0.01) ICC 0.39	N = 5539 0.17 (0.02 to 0.32)	N = 5533 -0.11 (-029 to 0.007)
Aggressive behaviour	N = 2584 1.70 (0.87)	N = 2294 1.66 (0.80)	N = 2585 1.58 (0.78)	N = 2319 1.69 (0.80)	N = 2627 1.68 (0.85)	N = 2295 1.67 (0.83)	N=4203 -0.13 (-0.23 to -0.04)* ICC 0.29	N=4019 0.01 (-0.09 to 0.13) ICC 0.22	N = 5539 -0.14 (-0.24 to -0.05)	N = 5533 0.10 (-0,10 to 0.12)
Internalising/ withdrawn	N = 2584 2.19 (0.85)	N = 2294 2.04 (0.85)	N = 2585 1.91 (0.79)	N = 2319 2.00 (0.83)	N = 2627 1.95 (0.75)	N = 2295 1.93 (0.81)	N=4203 -0.16 (-0.27 to -0.04)	N=4009 0.01 (-0.12 to 0.22)	N = 5539 -0.17 (-0.26 to -0.08)	N = 5533 0.02 (-0.11 to 0.14)

[9] *3-level (school, class, child) hierarchical model with random effects model and adjusting for baseline outcome; negative score indicates follow-up score in intervention group < control group

[10] *3-level (school, class, child) hierarchical model with random effects model and adjusting for baseline outcome, age, gender, ethnicity, qualify for free school meal, special education needs and for school level stratification variables i.e. size of school, proportion of free school meals and for school level stratification variables i.e. size of school, proportion of free school meals. Negative score indicates follow up score in intervention group < control group

Inattention-hyperactivity	N = 2589 0.67 (0.76)	N = 2310 0.60 (0.71)	N = 2588 0.52 (0.67)	N = 2318 0.56 (0.68)	N = 2624 0.57 (0.71)	N = 2299 0.53 (0.68)	ICC 0.28 N=4226 -0.06 (-0.13 to -0.02)*	ICC 0.28 N=4040 0.04 (-0.04 to 0.12)	N = 5539 -0.06 (-0.13 to 0.02)	N = 5533 0.05 (-0.03 to 0.13)
Impulsivity-hyperactivity	N = 2589 0.53 (0.69)	N = 2310 0.47 (0.62)	N = 2588 0.42 (0.61)	N = 2318 0.46 (0.63)	N = 2624 0.49 (0.67)	N = 2299 0.45 (0.65)	ICC 0.19 N=4226 -0.06 (-0.13 to -0.01)*	ICC 0.24 N=4040 0.03 (-0.04 to 0.11)	N = 5539 -0.08 (-0.14 to -0.01)	N = 5533 0.04 (-0.04 to 0.11)
Peer relations	N = 2580 1.71 (0.70)	N = 2317 1.65 (0.67)	N = 2591 1.54 (0.64)	N = 2311 1.61 (0.70)	N = 2596 1.68 (0.73)	N = 2284 1.59 (0.71)	ICC 0.21 N=4217 -0.12 (-0.22 to -0.02)*	ICC 0.25 N=4003 0.08 (-0.04 to 0.21)	N = 5539 -0.13 (-0.22 to -0.03)	N = 5533 0.08 (-0.04 to 0.21)
Relational aggression	N = 2583 1.48 (0.78)	N = 2315 1.46 (0.73)	N = 2590 1.49 (0.78)	N = 2310 1.58 (0.83)	N = 2572 1.61 (0.86)	N = 2286 1.56 (0.82)	ICC 0.33 N=4217 -0.08 (-0.20 to 0.01)	ICC 0.42 N=3998 0.05 (-0.07 to 0.16)	N = 5539 -0.10 (-0.21 to -0.001)	N = 5533 0.05 (-0.07 to 0.17)
Learning behaviours	N = 2581 1.57 (0.41)	N = 2292 1.61 (0.38)	N = 2589 1.65 (0.39)	N = 2284 1.61 (0.38)	N = 2592 1.60 (0.41)	N = 2277 1.60 (0.40)	ICC 0.19 N=4180 0.05 (0.003 to 0.10)*	ICC 0.20 N=3974 -0.01 (-0.07 to 0.05)	N = 5539 0.05 (0.01 to 0.10)	N = 5533 -0.005 (-0.06 to 0.05)
Academic performance	N = 2577 3.00 (1.33)	N = 2319 3.02 (1.36)	N = 2584 3.14 (1.34)	N = 2279 3.11 (1.36)	N = 2594 3.08 (1.33)	N = 2279 3.10 (1.31)	ICC 0.09 ICC 0.25 N = 4184 0.03 (-0.03 to 0.10)	ICC 0.09 N=3996 -0.03 (-0.10 to 0.05)	N = 5539 0.03 (-0.04 to 0.09)	N = 5533 -0.04 (-0.13 to 0.05)
							ICC 0.03	ICC 0.05		

Note. * indicates $p < 0.05$

Table 5

T-POT Composite Categories

T-POT Categories	Intervention (baseline) (sd)	Control (baseline) (sd)	Intervention (6 months post-baseline) (sd)	Control (6 months post-baseline) (sd)	Effect size of ANCOVA
Total teacher positive behaviours	89.70**	82.78**	105.50**	73.22**	.304
	(31.89)	(26.02)	(29.65)	(22.59)	
Total teacher negative behaviours	15.00	9.56	8.00	10.89	.044
	(10.18)	(7.81)	(6.13)	(8.68)	
Total teacher praise	27.20	22.44	29.00	24.89	.002
	(11.82)	(10.68)	(15.02)	(17.18)	
Class compliance	50.10	53.44	61.80	63.44	.009
	(19.43)	(18.34)	(15.18)	(19.11)	
Class non-compliance	1.10	3.56	.80	1.67	.126
	(1.66)	(4.59)	(1.87)	(1.73)	
Class negative behaviour to teacher	6.10*	6.67*	5.70*	5.60*	.307
	(3.90)	(8.35)	(3.59)	(5.77)	
Class prosocial behaviour	25.40	25.33	25.00	19.33	.074
	(8.19)	(11.92)	(6.72)	(13.75)	
Class off-task behaviour	7.00*	2.22*	1.80*	5.22*	.227
	(6.45)	(1.79)	(3.42)	(4.82)	
Class total negative behaviours	9.50	8.11	3.50	7.44	.150
	(6.83)	(10.03)	(3.63)	(6.33)	

Note. *p<.05, **p<.01

Table 6

Strengths & Difficulties Questionnaire – Total Difficulties Moderator Analyses[11]

Characteristics	Subgroup analysis – SDQ total score, Interaction term on between-group difference+, Mean (95% CI)*, P-value	
	12 months post-baseline adjusted for baseline (N=4699)	24 months post-baseline adjusted for baseline (N=4473)
Child level		
Gender	-0.21 (-0.22 to 0.63) P-value 0.34	-0.33 (-0.87 to 0.21) P-value 0.23
Age	-0.60 (-1.22 to 0.01) P-value 0.06	0.01 (-0.76 to 0.78) P-value 0.98
Ethnicity – white vs. all	-0.43 (-1.00 to 0.14) P-value 0.14	-0.45 (-1.16 to 0.26) P-value 0.22
Ethnicity – Pakistani vs. all	0.24 (-0.15 to 0.64) P-value 0.22	0.15 (-0.35 to 0.65) P-value 0.56
Qualifying for free school meals	0.49 (0.02 to 0.95) P-value 0.04*	0.32 (-0.27 to 0.91) P-value 0.29
Special education needs	0.17 (-0.44 to 0.78) P-value 0.59	-0.06 (-0.83 to 0.72) P-value 0.89
SDQ total score cut-off	0.08 (-0.47 to 0.63) P-value 0.77	0.24 (-0.45 to 0.94) P-value 0.49
SDQ conduct score cut-off	-0.18 (-0.77 to 0.40) P-value 0.53	0.16 (-0.57 to 0.90) P-value 0.66
SDQ emotional score cut-off	-1.15 (-1.88 to -0.42) P-value 0.002*	-2.33 (-3.26 to -1.41) P-value <0.0001*
SDQ pro-social score cut-off	-0.11 (-0.54 to 0.31) P-value 0.60	-0.09 (-0.62 to 0.44) P-value 0.75
School level		
School size	0.02 (-0.95 to 0.99) P-value 0.97	-0.67 (-1.90 to 0.55) P-value 0.28

[11] +Based on LOCF models (all N=4699)

*3-level (school, class, child) hierarchical model with random effects model and adjusting for baseline outcome, age, gender, ethnicity, qualifying for free school meals, special education needs.

Provide SEAL	0.00 (0.00 to 0.001) P-value 0.11	0.001 (0.00 to 0.002) P-value 0.03*
Other		
Cohort	0.48 (-0.41 to 1.38) P-value 0.29	0.54 (-0.58 to 1.67) P-value 0.34

Table 7
Programme Costs and Cost Per Child of Running PATHS Programme for 29 Schools for 2009-10

			Total cost (£)[1]
Non-recurring group set-up costs to train 101 teachers[2] / school			
Materials (programme kit)			29,472
Incentives, key rings etc			27,422
Trainer fees			29,090
Administration (no information available)			
Teacher time attending workshops (1.5 days x 101 teachers x £27.11/hr)			30,804
Supply cover for teachers attending training			9,864
Venue and accommodation			13,849
Non-recurring training cost total			**140,501**
Non-recurring training cost per teacher (above / by 101 teachers)			**1,391**
Non-recurring training cost per teacher annuitized over five years at 3.5% / yr			308
Non-recurring training cost per school (total cost/ 29 schools)			4,845
Non-recurring training cost per school annuitized over five years at 3.5% / yr			1,073
Recurring group running costs across 29 schools			
Project management (based on costs supplied by BCC)			Cost (£)
	Gross salary	WTE +time employed	Total / school
Pilot Project Manager	£54,325 / year	1 WTE for 3 months	13,581
Pilot Project Manager (Interim Consultant)	£68,821 / year	1 WTE for 4 months	22,940
Administrative support for PATHS	£13,287 / year	0.33 WTE for 1 year	4,429
Project management total across 29 schools			**40,950**
Project management sub-total for one school (above / by 29			**1,412**

schools)

Delivery of PATHS programme for 29 schools[3]	Mean Unit cost (£)	Number of units	
Two PATHS lessons plus extension lesson delivered by 101 teachers (based on mid Pt M4-6 qualified teacher salary £29,328[1])	£27.11/hr	1 hr /wk for 40 wks	109,524
Preparation time (based on estimates from four teachers)	£27.11/hr	0.75 hr / wk for 40 wks	82,143
PATHS consultant coaches (x3)			103,430
PATHS coaches travel expenses			152
Delivery of PATHS sub-total across 29 schools			**295,249**
Sub-total per school			**10,181**

Cost (£) of establishing, project managing and delivering PATHS in 29 schools over a 40-week school year, inclusive of training annuitised over five years at 3.5% / year	
Cost per school	**12,666**
Cost per child in Years One & Two (N=2640)	**139**
Cost (£) of project managing and delivering PATHS in 29 schools over 40-week school year	
Cost per school	**11,593**
Cost per child in Years One & Two receiving PATHS (N=2640)	**127**

[1] Hourly salary calculation based on annual salary divided by 38 weeks/37.5 hrs including 25% employers cost; 38 weeks allows for holiday, CPD and sickness.
[2] Based on costs supplied by BCC apart from costs related to teachers attending sessions
[3] Based on PATHS time estimates, national average salary, coaches costs from BCC, supplemented by four teachers