

Case study 1: An Evidence-based practice review report.

Theme: School/Setting Based Interventions for Social, Emotional and Mental Health.

Effectiveness of the Promoting Alternative Thinking Strategies (PATHS®) programme for developing social-emotional competence among primary school pupils aged 4-11 years.

Summary

Social-emotional learning (SEL) interventions aim explicitly to teach children skills to promote social-emotional competence (Panayiotou et al., 2020). There is an increasing evidence base for the effectiveness of SEL programmes across a wide range of outcomes for children, including their academic achievement, social-emotional competence and mental well-being (Durlak et al., 2011; Sklad et al., 2012). The Promoting Alternative Thinking Strategies (PATHS®) curriculum is an increasingly popular SEL intervention that has been used widely within the UK education system. The current review aims to evaluate the effectiveness of PATHS® in improving social-emotional competence among children aged 4-11 years old. Seven papers met the inclusion criteria for the current review and were evaluated using Gough's (2007) weight of evidence framework to critically appraise the evidence. A detailed analysis of included studies found that, overall, limited evidence exists to suggest that the intervention improved social-emotional competence among this age group. Potential implications for practice within the UK education system are considered and future research needs are discussed.

Introduction

Social and Emotional Learning

Social and emotional learning (SEL) is defined as “the processes through which all young people and adults acquire and apply the knowledge, skills, and attitudes to develop healthy identities, manage emotions, achieve personal and collective goals, feel and show empathy for others, establish and maintain supportive relationships, and make responsible and caring decisions” (Collaborative for Academic, Social and Emotional Learning [CASEL], 2021).

SEL involves the development of both intra- and inter-personal skills. Intra-personal skills relate to our internal abilities to manage emotions and behaviours, such as emotion-regulation and self-control. An ability to process, identify, label and tolerate our emotions has an important role in promoting mental health and well-being (Feldman-Barrett et al., 2001; Kabat-Zinn, 2003; Lischetzke & Eid, 2003). Inter-personal skills are those that enable us to relate to others, such as perspective-taking and social problem-solving which impact on relationships (Berry et al., 2016; Domitrovich et al., 2017).

Social-emotional competence is associated with success in several outcomes throughout childhood and into adulthood (Domitrovich et al., 2017). It enhances our ability to behave in a socially appropriate manner, reduces risk-taking behaviour, promotes the development of positive relationships and is also associated with academic progress (Epstein et al., 2000; Trentacosta & Fine, 2010). Longitudinal research demonstrates that when compared to cognitive ability, social-emotional competence in childhood is more strongly associated with well-being, health, lifestyle

and some labour market outcomes in adulthood including employment status (Goodman et al., 2015).

Social-emotional competence is thought to follow a developmental trajectory, shaped through social experiences within a child's environment (Humphrey et al., 2018). These social processes can be implicit or explicit and occur across a range of settings (Saarni, 1999).

Given the significant amount of time that children spend in school throughout their education, schools provide a unique opportunity to teach and develop social-emotional competence which promotes well-being and reduces the likelihood of negative long-term outcomes (Domitrovich et al., 2017). This can be supported through the use of school-based SEL interventions (Department of Health, 2013). SEL interventions explicitly teach children social and emotional skills that are important for success within the school environment and later life (Taylor et al., 2017). The implementation and evaluation of school-based SEL interventions has increased in recent years (Greenberg et al., 2003) and the growing role of schools in promoting social-emotional skills has been reflected in government legislation and recommendations (Department for Education & Department of Health & Social Care, 2018; Weare & Gray, 2003).

Meta-analyses have reported positive effects of SEL interventions among primary school children including improved social-emotional competence, self-image, mental health, academic progress and fewer antisocial behaviours (Durlak et al., 2011; Sklad et al., 2012). The current review aims to further explore the evidence base for

PATHS®, a SEL intervention that has become popular internationally and more recently within the UK.

Promoting Alternative Thinking Strategies (PATHS®)

PATHS® was originally designed for use among hearing impaired children in the USA to develop social-emotional competence (Greenberg & Kusche, 1993), and has since been modified for use in various special education and mainstream schools. PATHS® is a universal intervention delivered to all pupils. It aims to improve the inter- and intra-personal skills of children aged 4-11 years old. PATHS® focuses on five core components (PATHS® UK, 2018): self-awareness, self-management, responsible decision-making, relationship skills and social awareness (see Table 1 for details).

PATHS® was influenced by several psychological theories and approaches. One of these is the Affective, Behavioural, Cognitive, Dynamic model of development (Greenberg & Kusche, 1993). This model emphasises that social-emotional competence requires several skills, including affective, behavioural and cognitive skills. The PATHS® curriculum aims to support children and young people to develop skills in each area and learn to better integrate these skills, resulting in social-emotional competence that facilitates more adaptive functioning (Curtis & Norgate, 2007; Kusche & Greenberg, 1994).

PATHS® is typically delivered by class teachers trained by an external coach. Schools are then provided with a PATHS® curriculum pack which contains around 40 lessons per year group. Lessons are typically delivered twice a week and follow a common format. Lessons start with an introduction to the session followed by group activities or a story delivered by the class teacher and then reflection time (Humphrey et al., 2018). A 'spiral curriculum' is used in which lessons are sequenced to be

developmentally appropriate, topics are revisited on different occasions and new learning is linked with previous topics.

PATHS® was also influenced by the eco-behavioural systems framework, originally developed by Roger Barker (Barker, 1968; Barker & Gump, 1964; Schogen, 1989). The model highlights the way that the settings children find themselves in, such as the classroom, playground or home environment, each have identifiable differences and associated social rules that influence and inform behaviour (Kumar, 2009). It suggests that interventions will be most effective when different settings around the child provide environmental changes or cues that allow for the skills learned during the curriculum to be used and generalised (Curtis & Norgate, 2007). The generalisation of skills in PATHS® is promoted using school-wide strategies for the regulation and management of emotions, along with optional activities for the home environment.

Rationale and Relevance

Educational Psychologists (EPs) have an ethical responsibility to work competently. This includes ensuring that recommended interventions have a current and accurate evidence base (British Psychological Society, 2018). Given the growing popularity of PATHS® within the UK, it is vital that the effectiveness of this intervention is evaluated. Although many large-scale trials have evaluated PATHS®, there has yet been no review of the effectiveness of the intervention among primary school children that focuses on social-emotional competence.

Review Question

The current review therefore aims to address the following question: How effective is the Promoting Alternative THinking Strategies (PATHS®) programme for developing social-emotional competence among primary school pupils aged 4-11 years?

Table 1: Core Components of the PATHS® Intervention (Adapted from PATHS® UK, 2018)

Curriculum Component	Description
Self-awareness	<p>This key component teaches children emotional understanding using a developmental approach throughout each year group. Activities to help children develop their emotional understanding include the use of 'feeling faces' and 'feeling dictionaries'.</p> <p>Self-esteem is also developed through the use of regular praise and the giving of compliments to children using the 'pupil of the day' celebration.</p>
Self-management	<p>Techniques taught help pupils to learn to recognise and manage emotions, feelings and behaviours. Strategies include the 'turtle technique', 'three steps for calming down' and 'control signals'.</p>
Responsible decision-making	<p>This component aims to support pupils to understand the steps involved in social problem solving using the '11 steps to problem solving poster'.</p>

Relationship skills	Children are supported to consider the role of feelings in social relationships and how to solve any problems that arise within their existing friendships. Techniques such as 'co-operative learning roles' and 'fair play rules' are used to foster positive social relationships within the school environment.
Social awareness	As children develop their understanding of and their ability to manage their own emotions, they are taught how to recognise others' emotions and behaviours.

Critical Review of the Evidence Base

Literature Search

A systematic search of the literature was carried out on 27th December 2020 using the online databases: Psych INFO, ERIC and Web of Science. A combination of key-word searches and subject heading searches were performed and are detailed in Table 2. Searches were limited to peer reviewed journal articles.

Table 2: List of Search Terms and Subject Headings Used for Search

Database	Keyword Search	Subject Heading Search
Psych INFO	Promoting Alternative Thinking Strategies OR PATHS OR Social-emotional learning program* OR Social emotional learning curriculum	Social Emotional Learning
	Social adj3 develop* OR Social skill* OR Social behav* OR Social understand* OR Social competenc* OR Social-emotional competenc* OR Social-emotional develop*	Emotional Regulation
	Child* OR P?ediatric* OR School-age* OR Juvenile* OR Minor* OR Primary school* OR Elementary school* OR Young OR Youth	
ERIC	“Promoting Alternative Thinking Strategies” OR “PATHS” OR “Social-emotional learning program*” OR “Social emotional N2 learning curriculum”	
	“Social N3 develop*” OR “Social skill*” OR “Social behav*” OR “Social understand*” OR “Social competenc*” OR “Social-emotional competenc*” OR “Social-emotional develop*”	Emotional Intelligence OR Interpersonal Competence OR Perspective Taking

	<p>“Child*” OR “P?ediatric*” OR “School-age*” OR “Juvenile*” OR “Minor*” OR “Primary school*” OR “Elementary school*” OR “Young” OR “Youth”</p>	Children
Web of Science	<p>“Promoting Alternative Thinking Strategies” OR “PATHS” OR “Social-emotional learning program*” OR “Social emotional NEAR/2 learning curriculum”</p> <p>“Social NEAR/3 develop*” OR “Social skill*” OR “Social behav*” OR “Social understand*” OR “Social competenc*” OR “Social-emotional competenc*” OR “Social-emotional develop*”</p> <p>“Child*” OR “P\$ediatric*” OR “School-age*” OR “Juvenile*” OR “Minor*” OR “Primary school*” OR “Elementary school*” OR “Young” OR “Youth”</p>	

Screening of Articles

The initial search used a combination of database searches and ancestral searches, generating a total of 611 results. After removing duplicates (n=93), 518 titles and abstracts were identified and screened according to inclusion and exclusion criteria (see Table 3). 27 articles were selected for full text screening and 20 studies were excluded (see Appendix A for reasons for exclusion from the review). The 7 studies included in the review are listed in Table 4. Further details of each study and main findings are provided in Appendix B. Figure 1 provides an overview of the systematic search process.

Table 3: Inclusion and Exclusion Criteria for Current Literature Search

	Inclusion	Exclusion	Rationale
Type of Publication	Peer reviewed article/journal.	Any publication that has not been peer reviewed.	Peer reviewed journals have previously been reviewed and scrutinised to ensure methodological integrity. They are therefore considered to be of a high methodological quality.
Study Type	Intervention studies that use an experimental or quasi-experimental research design and provide original empirical data to evaluate the impact of the PATHS® curriculum on social-emotional competence.	Articles that do not use an experimental method or review findings from previously conducted studies and therefore do not include original empirical data.	The review aims to evaluate the effectiveness of the PATHS® intervention on social-emotional competence and therefore relies on original empirical research.
Publication Date	Studies with a publication date between 2000 and 27 th December 2020.	Any studies with a publication date prior to 2000.	Publication within the past 20 years was selected to ensure that all evidence included remains relevant to current educational practice and settings.

Age of Participants	Participants in the study are of English primary school age (4-11 years)	Participants that are either younger than 4 years old or older than 11 years of age.	The review is evaluating an intervention for primary school aged pupils.
Settings	The intervention is carried out in a mainstream school setting by education staff.	The intervention is delivered in settings that are not educational or by professionals not typically based in schools.	The review aims to evaluate the effectiveness of the PATHS® curriculum when delivered by school staff within educational settings.
Intervention	The PATHS® intervention is included in at least one or more of the intervention conditions.	The included intervention conditions do not use PATHS® as an intervention for social-emotional competence.	The review aims to evaluate the effectiveness of the PATHS® curriculum.
Outcomes	The study has at least one outcome measuring social and/or emotional competence.	The study does not include an outcome measure that measures social and/or emotional competence.	This review aims to evaluate the effectiveness of the PATHS® curriculum specifically on social-emotional competence.
Geographical Context and Language of Publication	Studies written in English from any geographical context.	Studies that are not published in English.	Only articles published in English can be read by the author which ensures that all included studies are fully understood and critically evaluated.

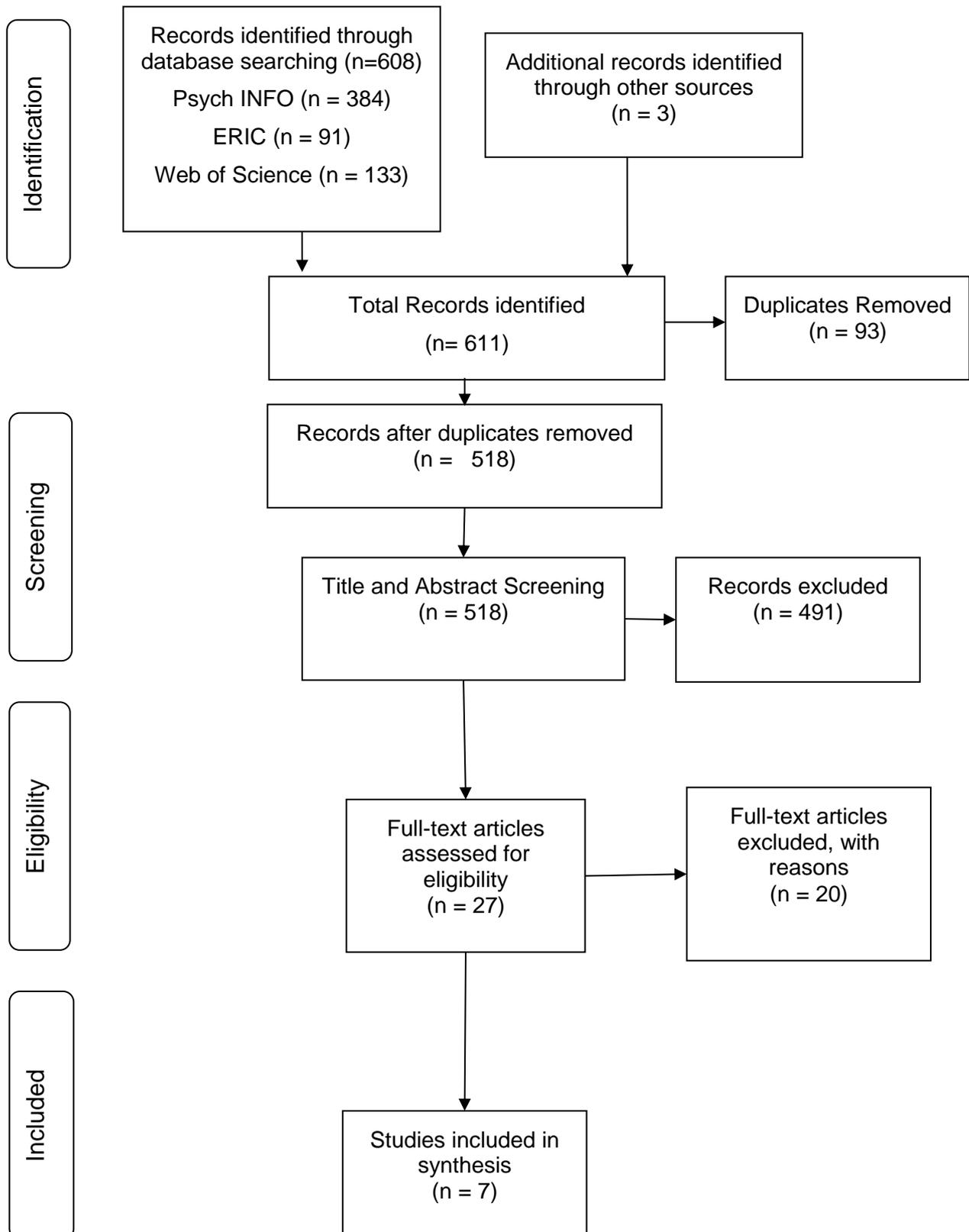


Figure 1: Flow diagram to depict the literature search and selection process based on PRISMA Statement Recommendations (Moher et al., 2010).

Table 4: Full References of Included Studies

Included Studies	
1	Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (2010). The effects of a multiyear universal social–emotional learning program: The role of student and school characteristics. <i>Journal of Consulting and Clinical Psychology, 78</i> (2), 156.
2	Novak, M., Mihic, J., Basic, J., & Nix, R (2017). PATHS in Croatia: A school-based randomised-controlled trial of a social and emotional learning curriculum. <i>International Journal of Psychology, 52</i> (2), 87–95.
3	Curtis, C., & Norgate, R. (2007). An evaluation of the Promoting Alternative Thinking Strategies curriculum at key stage 1. <i>Educational Psychology in Practice, 23</i> (1), 33–44.
4	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. <i>International Journal of Conflict and Violence, 6</i> (2), 260–272.
5	Humphrey, N., Hennessey, A., Lendrum, A., Wigelsworth, M., Turner, A., Panayiotou, M., Joyce, C., Pert, K., Stephens, E., Wo, L., Squires, G., Woods, K., Harrison, M., & Calam, R. (2018). The PATHS curriculum for promoting social and emotional well-being among children aged 7–9 years: a cluster RCT. <i>Public Health Research, 6</i> (10), 1–116.
6	Seifer, R., Gouley, K., Miller, A. L., & Zakriski, A. (2004). Implementation of the PATHS Curriculum in an Urban Elementary School. <i>Early Education and Development, 15</i> (4), 471–486.
7	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. <i>Journal of Clinical Child and Adolescent Psychology, 40</i> (5), 677–692.

Mapping the Field

The studies included in the review vary in terms of the research design employed, methods of evaluating social-emotional competence and the details of intervention implementation. However, all studies investigated the effectiveness of PATHS® on social-emotional competence among pupils aged 4-11 years.

Weight of Evidence

Gough's (2007) Weight of Evidence Framework was used to appraise the selected studies. This framework allows a systematic evaluation of research based on three key dimensions: methodological quality, methodological relevance, and topic relevance.

Weight of Evidence A (WoE A) considered the methodological quality of the research using an adapted version of Kratochwill's (2003) group intervention protocol. The protocol considers key methodological features including measurement, comparison groups, identifiable components, implementation fidelity and follow-up assessment. WoE B considered the methodological relevance of research and its appropriateness to the review question based on the recommendations of Petticrew and Robert's typologies of evidence (2003). WoE C evaluated the topic relevance of the research for the current review question (Gough, 2007). Appendices C-J provide further details of the coding protocols and criteria used for evaluation, along with details of any adaptations that were made to existing protocols.

The overall ratings for WoE A, B and C were given equal weightings and were totalled and divided by three to give an overall WoE D rating (Table

5). WoE D provides a summary of the overall strength of the evidence in answering the current review question.

Table 5: Overall Weight of Evidence D Ratings

Study	WoE A Methodological Quality	WoE B Methodological Relevance	WoE C Topic Relevance	WoE D Overall weight of evidence
Bierman et al. (2010)	1.2	2.5	2.5	2.1 Medium
Novak et al. (2017)	1.2	2.5	2.0	1.9 Medium
Curtis & Norgate (2007)	1.2	2.0	2.3	1.8 Medium
Little et al. (2012)	1.4	2.5	2.3	2.1 Medium
Humphrey et al. (2018)	1.6	3.0	3.0	2.5 High
Seifer et al. (2004)	0.6	1.8	1.8	1.4 Low
Malti et al. (2011)	2.0	2.5	2.5	2.3 Medium

Note: ≤ 1.4 constitutes a low score, 1.5-2.4 is a medium score and ≥ 2.5 is a high score.

Participants

A total of 14,841 participants from four countries (UK, USA, Switzerland and Croatia) were included in the review. The relevance of the country in which the study was conducted relative to the UK context was reflected in WoE C ratings. UK based studies were given the highest rating and OECD countries were rated more highly than non-OECD countries as they were likely to be most similar to the UK context.

Studies varied in terms of sample size. The smallest sample consisted of 150 participants (Seifer et al., 2004) whereas 5,218 participants were included in the largest sample (Humphrey et al., 2018). Participants' ages ranged from 4 (Little et al., 2012) to 11 years (Malti et al., 2011). Three studies

reported the gender of participants in the intervention group. The percentage of females ranged from 47% (Novak et al., 2017) to 50.1% (Humphrey et al., 2018).

Study Design

The research design of each study was considered using the WoE A and B ratings. A hierarchy of evidence developed by Petticrew and Roberts (2003), suggests that when addressing the 'effectiveness' of a group intervention the most appropriate research design is a Randomised Controlled Trial (RCT).

Five studies (Bierman et al., 2010; Humphrey et al., 2018; Little et al., 2012; Malti et al., 2011; Novak et al., 2017) used a cluster RCT and gained a high WoE B rating. Cluster RCTs randomise groups of participants (e.g., classes or schools) to study conditions rather than individuals. All the cluster RCTs used random allocation of participants to the intervention and control groups. In two studies (Humphrey et al., 2018; Little et al., 2012) the randomisation procedure was carried out independently of the research team. Employing an independent research team to randomise groups reduces the possibility of bias, therefore increasing the likelihood that the results of the study can be generalised to a wider population (Barker et al., 2015). One study (Malti et al., 2011) reported computerised random allocation of participants to control or intervention groups.

Two studies used a group experimental design (Curtis & Norgate, 2007; Seifer et al., 2004) and received a lower WoE B rating. Curtis and Norgate (2007) did not randomly allocate participants to control or intervention groups as the control condition in their study was due to receive the PATHS®

intervention the following academic year. Seifer et al. (2004) also used non-random allocation of participants to control and intervention groups. Both studies received a low WoE A rating because of these limitations.

Control groups varied between studies and were evaluated in WoE A and WoE B. Three studies (Curtis & Norgate, 2007; Novak et al., 2017; Seifer et al., 2004) reported a 'usual practice' control group. Two studies (Humphrey et al., 2018; Little et al., 2012) specified 'waitlist' control groups in which participants had access to existing social and emotional teaching within school. These studies all received a 'medium' rating. No studies received a 'high' rating as they did not include an 'active' comparison group that compared the effects of PATHS® with another commonly used SEL intervention. Two studies (Bierman et al., 2010; Malti et al., 2011) did not specify the control conditions used.

Implementation and Fidelity

All included studies implemented PATHS®. Novak et al. (2017) translated the curriculum to Croatian using a professional translator. It was checked for content and nuance by the authors, who spoke both English and Croatian, and are experienced within the field of developmental psychology. PATHS® was adapted by Malti et al. (2011) to make it culturally appropriate for the Swiss school system. The authors reported that adapted materials were tested in a pilot study to ensure their appropriateness.

Intervention duration varied between studies, along with the frequency and duration of PATHS® lessons (see Table 6 for details). Interventions lasted between one academic year (Curtis & Norgate, 2007; Malti et al., 2011; Novak et al., 2017) and three years (Bierman et al., 2010; Humphrey et al., 2018).

The total number of PATHS® lessons delivered per year ranged from 43 (Humphrey et al., 2018) to 63 (Novak et al., 2017). PATHS® lessons were delivered between 1-3 times per week and ranged from 20-30 minutes (Bierman et al., 2010) to 67 minutes (Malti et al., 2011). In all studies, the class teacher delivered the intervention which promotes the feasibility of the intervention and generalisation of skills learned.

Implementation fidelity is reflected in WoE C ratings. Four studies (Bierman et al., 2010; Humphrey et al., 2018; Malti et al., 2011; Novak et al., 2017) used independent observers to provide feedback and ongoing support for teachers. These studies were given a 'high' WoE C rating as the support received is likely to influence the quality of intervention delivery. Two studies (Curtis & Norgate, 2007; Little et al., 2012) received a 'medium' rating because although they relied upon self-report measures rather than independent observations, the class teachers did have technical support from trained PATHS® coaches. Finally, Seifer et al. (2004) received a 'low' rating as only self-report measures were used.

Table 6: Summary of PATHS Implementation Across Included Studies

Study	Total Duration of Intervention	Number of PATHS® Lessons	PATHS® Lesson Duration	Frequency of PATHS® lessons	Interventionist
Bierman et al. (2010)	3 academic years	1 st Grade: 57 2 nd Grade: 46 3 rd Grade: 48	20-30 minutes	2-3 per week	Class Teacher
Novak et al. (2017)	1 academic year	63	Not Reported	2 per week	Class Teacher
Curtis & Norgate (2007)	1 academic year	Not Reported	Not Reported	Not Reported	Class Teacher
Little et al. (2012)	2 academic years	Not Reported	60 minutes	1 per week	Class Teacher
Humphrey et al. (2018)	3 academic years	Year 3 = 43 Year 4 = 57 Year 5 = 55	30-40 minutes	2 per week	Class Teacher
Malti et al. (2011)	1 academic year	46	Mean duration of 67 minutes	Mean frequency of 2.4 per week	Class Teacher
Seifer et al. (2004)	2 academic years	Not Reported	30-45 minutes	3 per week	Class Teacher

Measures

Studies varied in the measures used to assess social-emotional competence. Some focused on overall measures of social-emotional competence (Bierman et al., 2010; Little et al., 2012; Seifer et al., 2004). Bierman et al. (2010) used the Teacher Observation of Classroom Adaptation-Revised (TOCA-R; Werthamer-Larsson et al., 1991) and the Social Health Profile (SHP; CPPRG, 1998). Little et al. (2012) used the PATHS Teacher Rating Survey (PTRS) which they specified was a composite measure of seven scales. Seifer et al. (2004) used the Psychological Impairment Rating Scale (PIRS; Gouley et al., 2003).

Other studies focused on specific skills such as prosocial behaviour. Novak et al. (2017) used the Social Competence Scale (SCS; Fast Track Project, n.d.), Curtis and Norgate (2007) and Humphrey et al. (2018) used the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) and Malti et al. (2018) used the Social Behaviour Questionnaire (SBQ; Tremblay et al., 1991).

Little et al. (2012) and Novak et al. (2017) also investigated the effectiveness of PATHS® on Emotional Regulation. Novak et al. (2017) assessed this using the SCS and Little et al. (2012) assessed emotional regulation using the PTRS as described above.

The quality of measures used was evaluated using the WoE A framework. Two studies received a 'high' rating (Humphrey et al., 2018; Malti et al., 2011) because the measures had evidenced high validity and reliability and relied on multiple sources and measures to collect data, enabling triangulation of findings.

Four studies received a ‘medium’ rating for measures (Curtis & Norgate, 2007; Bierman et al., 2010; Little et al., 2012; Novak et al., 2017). Some of these used measures with high validity and reliability (such as Little et al., 2012) but did not achieve a ‘high’ rating as they failed to use measures from different sources. Seifer et al. (2004) obtained a ‘low’ rating as no details were provided regarding the reliability of measures. A lack of reliability data limits the conclusions that can be drawn from the research as differences between groups could be caused by unreliable assessment tools rather than the intervention.

Only one study (Humphrey et al., 2018) carried out follow-up assessments on completion of the intervention (at 12 and 24 months post-intervention). This study received a higher WoE A rating for the follow-up section of the protocol. However, the follow-up was limited as approximately two thirds of participants were not included, making it difficult to know whether the results of the intervention were maintained over time across the whole sample.

Findings

Tables 7-9 provide a summary of the effect sizes for all studies in the current review, calculated using Cohen’s *d*. The interpretation of the effect size value relied upon Cohen’s (1992) descriptions (see Table 10).

Table 10: Cohen’s *d* qualitative descriptions

Effect Size	Description
0.2	Small
0.5	Medium
0.8	Large

Table 7: Summary of Effect Sizes for Included Studies: Pro-social Behaviour Outcomes

Study	Sample Size	Outcome Measures for Prosocial Behaviour	Effect Size Cohen's <i>d</i>	Effect Size Descriptor	Significance Value	Overall WoE D Rating
Novak et al. (2017)	N=568 Intervention N = 280 Control N = 288	Social Competence Scale: Prosocial Behaviour Subtest.	0.16	Negligible	$p > 0.05$	1.9 (Medium)
Curtis & Norgate (2007)	N = 287 Intervention N = 114 Control N = 173	Strengths and Difficulties Questionnaire: Prosocial Behaviour Subtest.	0.91	Large	$p < 0.001^*$	1.8 (Medium)
Little et al. (2012)	N = 5953 Intervention classes = 102 Control classes = 94	Strengths and Difficulties Questionnaire: Prosocial Behaviour Subtest	0.02 -0.02	Negligible Negligible	$p > 0.05$ $p > 0.05$	2.1 (Medium)
Humphrey et al. (2018)	N = 5,218 Intervention N= 2294 Control N = 2106	PATHS Teacher Rating Survey: Prosocial Scale Strengths and Difficulties Questionnaire: Prosocial Behaviour Subtest	0.07	Negligible	$p > 0.05$	2.5 (High)
Malti et al. (2011)	N= 1,675 Intervention N = 360 Control N = 356	Social Behaviour Questionnaire: Prosocial Behaviour (Teacher report)	0.14	Negligible	$p > 0.05$	2.3 (Medium)

Table 8: Summary of Effect Sizes for Included Studies: Emotion Regulation Outcomes

Study	Sample Size	Outcome Measures for Emotion Regulation	Effect Size Cohen's <i>d</i>	Effect Size Descriptor	Significance Value	Overall WoE D Rating
Novak et al. (2017)	N =568 Intervention N = 280 Control N = 288	Social Competence Scale: Emotion Regulation Subtest	0.18	Negligible	$p < 0.10$	1.9 (Medium)
Little et al. (2012)	N = 4019 Intervention classes = 102 Control classes = 94	PATHS Teacher Rating Survey: Emotional Regulation Subtest	-0.02	Negligible	$p > 0.05$	2.1 (Medium)

Table 9: Summary of Effect Sizes for Included Studies: Social-Emotional Competence Outcomes

Study	Sample Size	Outcome Measures for Social-Emotional Competence	Effect Size Cohen's <i>d</i>	Effect Size Descriptor	Significance Value	Overall WoE D Rating
Little et al. (2012)	N = 4019 Intervention classes = 102 Control classes = 94	PATHS Teacher Rating Survey: Social Competence Scale	-0.01	Negligible	$p > 0.05$	2.1 (Medium)
Seifer et al. (2004)	N =150 Intervention N = 62 Control N = 88	Psychological Impairment Rating Scale: Social and Emotional Composite	0.38	Small	$p < 0.05^*$	1.4 (Low)
Bierman et al. (2010)	N = 2,937 Intervention Classes = 190 Control Classes = 180	Teacher Observation of Classroom Adaptation-Revised and Social Health Profile.	0.34	Small	$p < 0.001^*$	2.1 (Medium)

Novak et al. (2017) found no significant effect of PATHS® on prosocial behaviour. They reported a small improvement in emotion regulation in the intervention group compared to the control group, but this was only significant at $p < 0.10$ rather than $p < 0.05$. The study received a WoE D rating of 'medium' suggesting that the findings should be given due weight in evaluating the efficacy of PATHS®.

Curtis and Norgate (2007) found a significant difference between intervention and control groups on the SDQ prosocial behaviour scale. The intervention group demonstrated a higher level of prosocial behaviour than controls post-test. No effect size was reported in the original paper so was calculated using the results of the ANOVA interaction and an online effect size calculator (Wilson, n.d.). The effect size obtained was large ($d = 0.91$), but this should be interpreted cautiously as the study received a 'low' WoE A due to methodological limitations. Significant group differences were reported on pre-test measures in this study. Participants in the intervention group had higher levels of emotional symptoms, hyperactivity and conduct prior to the intervention when compared to the control group and a lower baseline level of prosocial behaviour. Thus, the effects of the intervention were not measured between comparable groups, significantly weakening the impact of the study's results.

Little et al. (2012) found no significant differences between control and intervention groups on prosocial behaviour, emotion regulation or overall social-emotional competence. As no effect sizes were reported, the mean difference between intervention and control group and 95% confidence intervals were used to calculate Cohen's d using the online calculator (Wilson,

n.d.). Although this study received a 'low' WoE A rating, its 'high' rating for its methodological relevance and 'medium' rating for topic relevance suggests that it provides useful and relevant evidence for the current review.

Humphrey et al. (2018) achieved the highest WoE rating of all studies in the review. No significant differences between intervention and control groups were found on the prosocial scale of the SDQ. Its findings are considered important, given that this study was of 'medium' methodological quality and highly relevant to the current review question.

Malti et al. (2011) reported no significant effect of PATHS® on prosocial behaviour as measured by the Social Behaviour Questionnaire (SBQ) but the results did suggest that prosocial behaviour changed in the expected direction between baseline and post-test measures. The analysis also controlled for a variety of child- and family-level factors. Increases in teacher-reported prosocial behaviour were influenced by the child being female, baseline levels of prosocial behaviour, Swiss background and low socio-economic status. This highlights the potential bias inherent in teacher-reported outcomes and the value of multiple sources of outcome data when considering the effectiveness of an intervention delivered by teachers.

Seifer et al. (2004) reported a small effect ($d = 0.38$, $p < 0.05$) of PATHS® on improved social-emotional competence. However, this study received a 'low' overall WoE D rating due to limitations associated with its methodological quality and topic relevance for the current review. For example, the authors used measures with no evidence of reliability and did not report a baseline measure. The findings therefore need to be interpreted with caution.

Finally, Bierman et al. (2010) reported a small but significant effect ($d = 0.34$, $p < 0.001$) of the PATHS® intervention on a measure of social-emotional competence. The study received 'high' ratings for its methodological relevance and relevance to the current review and therefore this evidence should be given due weight. It is, however, important to note that no baseline measures were completed pre-intervention.

Conclusions and Recommendations

This review evaluated the effectiveness of PATHS® for improving the social-emotional competence of pupils aged 4-11 years. There was some evidence that PATHS® achieves this. However, only two studies (Bierman et al., 2010; Seifer et al., 2004) reported significant findings with small effect sizes and one of these studies (Seifer et al., 2004) received a 'low' WoE D rating.

Regarding prosocial behaviour, one study found the PATHS® intervention to have a significant impact with a large effect size (Curtis & Norgate, 2007). The study received an overall WoE rating of 'medium' and was highly relevant to the current review question. However, it received a 'low' rating for its methodological quality and rigour which highlights the need to interpret the findings with caution. Finally, the review found no evidence that the intervention resulted in significant improvements in primary school pupils' emotion-regulation.

The review's conclusions could be considered unexpected given that the PATHS® curriculum was designed to promote social and emotional skills (Domitrovich et al., 2007) and is widely regarded as having an international evidence base (Humphrey et al., 2018). The implications of this review for practice and future research therefore require careful consideration.

Firstly, it has been argued that conventional methods of evaluating universal interventions do not fully capture the range of effects that the interventions produce. Greenberg and Abenavoli (2017) argue that a focus on the main effects of an intervention could miss clinically meaningful effects by failing to account for the heterogeneity that exists within universal samples. Universal interventions such as PATHS® may result in significant improvements in social-emotional competence for 'high risk' groups, whilst teaching basic social skills to children who have already mastered them will have very little effect (Engel et al., 2013). Approaches that consider the variation within a population could therefore shed light on important moderators of the interventions on certain subgroups of the population (Greenberg & Abenavoli, 2017). An interesting area for further research could therefore involve person-oriented techniques in the analysis of data to establish whether there are certain sub-groups that benefit most from PATHS® that would be missed by focusing on overall main effects of the intervention (Nix et al., 2016; Thase et al., 2011).

Another important factor relates to the implications of the PATHS® curriculum being implemented outside the country in which it was originally developed. In the current review, two studies (Bierman et al., 2010; Seifer et al., 2004) which demonstrated a significant effect of the PATHS® curriculum on social-emotional competence were conducted in the USA, where the intervention was developed. The remaining studies were conducted in other countries (Croatia, Switzerland and the UK). Cultural transferability is a particularly important consideration when evaluating the use of SEL programmes as social and emotional skills are strongly influenced by existing

cultural values and societal expectations (Wigelsworth et al., 2016). It could be that adaptations made to enable the PATHS® curriculum to be implemented across cultures influenced critical aspects of the intervention and impacted its efficacy (Wigelsworth et al., 2016). The impact of cultural transferability has been observed with other interventions (e.g. anti-bullying programmes) that were considered 'successful' in the USA but had mixed results when implemented within the UK (Ttofi et al., 2008). Considering how PATHS® can be modified to suit the UK education systems is therefore another area for future research. EPs could play an important role in such research, drawing on their expertise in child development, the education system and societal values and expectations.

Overall, the current review does not provide strong evidence for the effectiveness of the PATHS® intervention in promoting social-emotional competence among pupils aged 4-11 years. In addition, the lack of follow-up data across the included studies makes it difficult to evaluate whether the intervention had any long-term preventative effects that were not yet apparent at the end of the intervention. EPs recommending this intervention to schools should therefore be aware of the need for further research into the long-term preventative impacts of PATHS®; the possible impact of culture and context on the intervention; and meaningful ways of evaluating whether the intervention is beneficial for certain subgroups of the general UK school population.

References

- Barker, C., Pistrang, N., & Elliott, R. (2015). *Research methods in clinical psychology: An introduction for students and practitioners*. John Wiley & Sons.
- Barker, R. G. (1968). *Ecological Psychology*. Stanford University Press.
- Barker, R. G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior*. Stanford University Press.
- Berry, V., Axford, N., Blower, S., Taylor, R. S., Edwards, R. T., Tobin, K., Jones, C., & Bywater, T. (2016). The Effectiveness and Micro-costing Analysis of a Universal, School-Based, Social–Emotional Learning Programme in the UK: A Cluster-Randomised Controlled Trial. *School Mental Health, 8*(2), 238–256. <https://doi.org/10.1007/s12310-015-9160-1>
- Bierman, K., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., Pinderhughes, E., & Conduct Problems Prevention Research Group. (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. <https://doi.org/10.1037/a0018607>
- British Psychological Society. (2018). *Code of Ethics and Conduct*. BPS. Retrieved February 6, 2021, from <https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct>
- CASEL. (2021). *Collaborative for Academic, Social and Emotional Learning: What is SEL?* Retrieved February 6, 2021, from <https://casel.org/what-is->

sel/

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.

Conduct Problems Prevention Research Group. (1998). *Technical report for the Social Health Profile*. Retrieved from <http://sanford.duke.edu/centers/child/fasttrack/techrept/s/shs/shs3tech.pdf>

Curtis, C., & Norgate, R. (2007). An evaluation of the promoting alternative thinking strategies curriculum at key stage. *Educational Psychology in Practice*, 23(1), 33–44. <https://doi.org/10.1080/02667360601154717>

Department for Education, & Department of Health & Social Care. (2018). *Government response to the consultation on “Transforming children and young people’s mental health provision: a green paper” and next steps*. Retrieved February 6, 2021 from <https://www.gov.uk/government/consultations/transforming-children-and-young-peoples-mental-health-provision-a-green-paper>

Department of Health. (2013). Annual Report of the Chief Medical Officer 2012: “Our Children Deserve Better: Prevention Pays.” In *Annual Report of the Chief Medical Officer 2012*.

Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children’s social and emotional competence: A randomized trial of the preschool “PATHS” curriculum. *The Journal of Primary Prevention*, 28(2), 67–91. [https://doi.org/http://dx.doi.org/10.1007/s10935-007-0081-](https://doi.org/http://dx.doi.org/10.1007/s10935-007-0081-0)

0

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017).

Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development*, 88(2), 408–416.

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.

<https://doi.org/http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x>

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching Students What They Already Know? The (Mis)Alignment Between Mathematics Instructional Content and Student Knowledge in Kindergarten. *Educational Evaluation and Policy Analysis*.

<https://doi.org/10.3102/0162373712461850>

Epstein, J. A., Griffin, K. W., & Botvin, G. J. (2000). Competence skills help deter smoking among inner city adolescents. *Tobacco Control*, 9(1), 33–39.

Feldman-Barrett, L., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, 15(6), 713–724.

<https://doi.org/10.1080/02699930143000239>

Goodman R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, 38, 581–6.

<https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>

- Goodman, A., Joshi, H., Nasim, B., & Tyler, C. (2015). Social and emotional skills in childhood and their long-term effects on adult life. *London: Institute of Education.*
- Gough, D. (2007). Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213–228.
- Gouley, K. K., Eguia, M., Seifer, R., Dickstein, S., & Sameroff, A. J. (2003). *The psychological impairment rating scale: A brief measure of child social-emotional competence.* Unpublished manuscript, Bradley Hospital, Brown University Medical School.
- Greenberg, M. T., & Abenavoli, R. (2017). Universal Interventions: Fully Exploring Their Impacts and Potential to Produce Population-Level Impacts. *Journal of Research on Educational Effectiveness, 10*(1), 40–67. <https://doi.org/10.1080/19345747.2016.1246632>
- Greenberg, M. T., & Kusche, C. A. (1993). *Promoting Social and Emotional Development in Deaf Children. The PATHS Project.*
<http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=eric&AN=ED369237&site=ehost-live&scope=site&custid=s8454451>
- Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing school-based prevention and youth development through coordinated social, emotional, and academic learning. *American Psychologist, 58*(6–7), 466.
- Humphrey, N., Hennessey, A., Lendrum, A., Wigelsworth, M., Turner, A., Panayiotou, M., Joyce, C., Pert, K., Stephens, E., Wo, L., Squires, G.,

Woods, K., Harrison, M., & Calam, R. (2018). The PATHS curriculum for promoting social and emotional well-being among children aged 7–9 years: a cluster RCT. *Public Health Research*, 6(10), 1–116.

<https://doi.org/10.3310/phr06100>

Kabat-Zinn, J. (2003). Mindfulness-Based Interventions in Context: Past, Present, and Future. *Clinical Psychology: Science and Practice*, 10(2), 144–156. <https://doi.org/10.1093/clipsy/bpg016>

Kratochwill, T. R. (2003). *Task Force on Evidence Based Interventions in School Psychology*. American Psychological Association.

Kumar, T. K. (2009). Understanding the education system: An eco-behavioural approach. *Economic and Political Weekly*, 61–69.

Kusche, C. A., & Greenberg, M. T. (1994). *The PATHS Curriculum: Promoting alternative thinking strategies*. Developmental Research and Programs.

Lischetzke, T., & Eid, M. (2003). Is attention to feelings beneficial or detrimental to affective well-being? Mood regulation as a moderator variable. *Emotion*, 3(4), 361.

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.

<https://doi.org/10.4119/UNIBI/ijcv.263>

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.

<https://doi.org/10.1080/15374416.2011.597084>

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International Journal of Surgery*, 8(5), 336–341.

Nix, R. L., Bierman, K. L., Heinrichs, B. S., Gest, S. D., Welsh, J. A., & Domitrovich, C. E. (2016). The randomized controlled trial of Head Start REDI: Sustained effects on developmental trajectories of social-emotional functioning. *Journal of Consulting and Clinical Psychology*, 84(4), 310–322. <https://doi.org/http://dx.doi.org/10.1037/a0039937>

Novak, M., Mihić, J., Bašić, J., & Nix, R. L. (2017). PATHS in Croatia: A school-based randomised-controlled trial of a social and emotional learning curriculum. *International Journal of Psychology*, 52(2), 87–95. <https://doi.org/10.1002/ijop.12262>

Panayiotou, M., Humphrey, N., & Hennessey, A. (2020). Implementation matters: Using complier average causal effect estimation to determine the impact of the Promoting Alternative Thinking Strategies (PATHS) curriculum on children’s quality of life. *Journal of Educational Psychology*, 112(2), 236.

PATHS® UK (2018). *The PATHS® Programme for Schools (UK Version) Findings from the implementation of the PATHS® Programme for Schools (UK Version) in N. Ireland (2014-2018)*. Retrieved February 6, 2021 from:

http://www.pathseducation.co.uk/pdfs/public/PATHS_In_Focus_Report_%28October_2018%29_E_Version.pdf

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology & Community Health, 57*(7), 527–529. <https://doi.org/10.1136/jech.57.7.527>

Saarni, C. (1999). *The development of emotional competence*. Guilford press.

Schogen, P. (1989). *Behaviour Settings*. Stanford University Press.

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.

https://doi.org/10.1207/s15566935eed1504_6

Sklad, M., Diekstra, R., Ritter, M., Ben, J., & Gravesteyn, C. (2012). Effectiveness of school-based universal social, emotional, and behavioral programs: Do they enhance students' development in the area of skill, behavior, and adjustment? *Psychology in the Schools, 49*(9), 892–909. <https://doi.org/10.1002/pits.21641>

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. *Child Development, 88*(4), 1156–1171.

<https://doi.org/10.1111/cdev.12864>

Thase, M. E., Larsen, K. G., & Kennedy, S. H. (2011). Assessing the 'true' effect of active antidepressant therapy v. placebo in major depressive

disorder: use of a mixture model. *The British Journal of Psychiatry*, 199(6), 501–507.

Tremblay, R. E., Loeber, R., Gagnon, C., Charlebois, P., Larivee, S., & LeBlanc, M. (1991). Disruptive boys with stable and unstable high fighting behavior patterns during junior elementary school. *Journal of abnormal child psychology*, 19(3), 285-300.

Trentacosta, C. J., & Fine, S. E. (2010). Emotion knowledge, social competence, and behavior problems in childhood and adolescence: A meta-analytic review. *Social Development*, 19(1), 1–29.

Ttofi, M. M., Farrington, D. P., & Baldry, A. C. (2008). *Effectiveness of programmes to reduce school bullying: A systematic review*. Swedish Council for Crime Prevention, Information and publications.

Weare, K., & Gray, G. (2003). *What works in developing children's emotional and social competence and wellbeing?* DfES Publications.

Wigelsworth, M., Lendrum, A., Oldfield, J., Scott, A., ten Bokkel, I., Tate, K., & Emery, C. (2016). The impact of trial stage, developer involvement and international transferability on universal social and emotional learning programme outcomes: a meta-analysis. *Cambridge Journal of Education*, 46(3), 347–376.

<https://doi.org/10.1080/0305764X.2016.1195791>

Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American journal of community psychology*, 19(4), 585-602.

Wilson, D. B. (n.d.). *Practical Meta-Analysis Effect Size Calculator* [Online calculator]. Retrieved February 6, 2021, from

<https://www.campbellcollaboration.org/research-resources/effect-size-calculator.html>

Appendix A: Excluded Studies Following Full Text Screening

Table 1: Studies Excluded at Full Text Screening

Reference	Reason for Exclusion	Inclusion/Exclusion Criteria Number
Schonfeld, D. J., Adams, R. E., Fredstrom, B. K., Weissberg, R. P., Gilman, R., Voyce, C., ... & Speese-Linehan, D. (2015). Cluster-randomized trial demonstrating impact on academic achievement of elementary social-emotional learning. <i>School Psychology Quarterly</i> , 30(3), 406.	The study did not contain at least one outcome measuring social and/or emotional competence.	7
Raimundo, R., Marques-Pinto, A., & Lima, M. L. (2013). The effects of a social-emotional learning program on elementary school children: The role of pupil's characteristics. <i>Psychology in the Schools</i> , 50(2), 165-180.	The included intervention conditions do not use PATHS as an intervention for social emotional competence.	6
Crean, H. F., & Johnson, D. B. (2013). Promoting alternative thinking strategies (PATHS) and elementary school aged children's aggression: Results from a cluster randomized trial. <i>American Journal of Community Psychology</i> , 52(1-2), 56-72.	The study does not include an outcome that measures social and/or emotional competence.	7
Kam, C. M., Wong, L. W. L., & Fung, K. M. S. (2011). Promoting social-emotional learning in	Article did not use an experimental method as did not contain a control	2

Chinese schools: A feasibility study of PATHS implementation in Hong Kong. <i>The International Journal of Emotional Education</i> , 3(1), 30-47.	group with which to compare the effects of the intervention.	
Johannes, E. M. (1969). Effects of PATHS (c) after school program on children's social environment and behaviour.	Not a peer reviewed journal.	1
Van Schoiack, L. (2001). Promoting social-emotional competence: Effects of a social-emotional learning program and corresponding teaching practices in the schools.	Not a peer reviewed journal.	1
Dunn, M. (2019). The Impact of a Social Emotional Learning Curriculum on the Social-Emotional Competence of Elementary-Age Students.	Not a peer reviewed journal.	1
Averdijk, M., Zirk-Sadowski, J., Ribeaud, D., & Eisner, M. (2016). Long-term effects of two childhood psychosocial interventions on adolescent delinquency, substance use, and antisocial behavior: a cluster randomized controlled trial. <i>Journal of Experimental Criminology</i> , 12(1), 21-47.	The data set included in the paper is not an original dataset as it re-analyses the data from Malti et al., (2011) research.	2
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	Primary outcome was not investigating the effect of the PATHS intervention on social and/or emotional competence but	7

PATHS curriculum. <i>Prevention science</i> , 7(1), 91-102.	the role of neurocognition in mediating behavioural outcomes.	
Humphrey, N., Barlow, A., & Lendrum, A. (2018). Quality matters: Implementation moderates student outcomes in the PATHS curriculum. <i>Prevention Science</i> , 19(2), 197-208.	Primary outcome was not investigating the effect of the PATHS intervention on social and/or emotional competence but the mediating role of implementation on student outcomes.	7
Kelly, B., Longbottom, J., Potts, F., & Williamson, J. (2004). Applying emotional intelligence: Exploring the promoting alternative thinking strategies curriculum. <i>Educational Psychology in Practice</i> , 20(3), 221-240.	Article did not use an experimental design as did not contain a control group with which to compare the effects of the intervention.	2
Doikou-Avliidou, M., & Dadatsi, K. (2013). Enhancing social integration of immigrant pupils at risk for social, emotional and/or behavioural difficulties: the outcomes of a small-scale social-emotional learning programme. <i>Emotional and Behavioural Difficulties</i> , 18(1), 3-23.	The study did not use PATHS as the intervention.	6
Mattera, S., Lloyd, C. M., Fishman, M., & Bangser, M. (2013). A first look at the Head Start CARES Demonstration: Large-scale implementation of programs to improve children's social-emotional competence. <i>OPRE report</i> , 47.	The study included children younger than 4 years of age (Pre-school children).	4

<p>Panayiotou, M., Humphrey, N., & Hennessey, A. (2020). Implementation matters: Using complier average causal effect estimation to determine the impact of the Promoting Alternative Thinking Strategies (PATHS) curriculum on children's quality of life. <i>Journal of Educational Psychology</i>, 112(2), 236.</p>	<p>The study's primary outcome is a measure of quality of life rather than social and/or emotional competence.</p>	<p>7</p>
<p>Goossens, F., Gooren, E., de Castro, B. O., Van Overveld, K., Buijs, G., Monshouwer, K., ... & Paulussen, T. (2012). Implementation of PATHS through dutch municipal health services: A quasi-experiment. <i>International Journal of Conflict and Violence (IJCV)</i>, 6(2), 234-248.</p>	<p>The intervention was not delivered in an educational setting.</p>	<p>5</p>
<p>Malti, T., Ribeaud, D., & Eisner, M. (2012). Effectiveness of a universal school-based social competence program: The role of child characteristics and economic factors. <i>International Journal of Conflict and Violence (IJCV)</i>, 6(2), 249-259.</p>	<p>This article used the same data set reported in Malti et al., (2011) paper and therefore does not contain original empirical data.</p>	<p>2</p>
<p>Bardon, L. A., Dona, D. P., & Symons, F. J. (2008). Extending classwide social skills interventions to at-risk minority students: A preliminary application of randomization tests combined with single-subject design methodology. <i>Behavioral Disorders</i>, 33(3), 141-152.</p>	<p>The study does not include an outcome measure that measures social and/or emotional competence.</p>	<p>7</p>

<p>Calhoun, B., Williams, J., Greenberg, M., Domitrovich, C., Russell, M. A., & Fishbein, D. H. (2020). Social Emotional Learning Program Boosts Early Social and Behavioral Skills in Low-Income Urban Children. <i>Frontiers in Psychology, 11</i>.</p>	<p>The data reported in the study evaluated the long-term outcomes of the Pre-school PATHS curriculum and therefore when receiving the intervention the children in this study were not within the specified age range.</p>	<p>4</p>
<p>Humphrey, N., Barlow, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., ... & Turner, A. (2016). A cluster randomized controlled trial of the Promoting Alternative Thinking Strategies (PATHS) curriculum. <i>Journal of School Psychology, 58</i>, 73-89.</p>	<p>This article was excluded as a more updated paper was available using the same original data set.</p>	<p>2</p>
<p>Humphrey, N., Barlow, A., Wigelsworth, M., Lendrum, A., Pert, K., Joyce, C., ... & Humphrey, N. (2015). Promoting Alternative Thinking Strategies (PATHS): Evaluation Report and Executive Summary. <i>Education Endowment Foundation</i>.</p>	<p>This article was excluded as a more updated paper was available using the same original data set.</p>	<p>2</p>

Appendix B: Summary of Studies Included in Review

Table 2: Mapping the Field Table

Author, Date and Title	Study Design	Country and Setting	Participants	Intervention Details	Outcome Measures	Main Findings
<p>Author: Bierman et al. (2010)</p> <p>Title: The Effects of a Multiyear Universal Social-Emotional Learning Program: The Role of Student and School Characteristics.</p>	<p>Design: Cluster Randomised Controlled Trial.</p> <p>Groups: Random Allocation.</p> <p>Intervention: 190 classrooms</p> <p>Control: 180 classrooms</p>	<p>Country: USA</p> <p>Setting: Classes from Elementary Schools across three different sites (Tennessee, Washington & Pennsylvania)</p>	<p>Sample size: 2,937</p> <p>Age: 6-9 years (Grade 1-3 in Elementary School)</p> <p>Gender: Not Reported</p>	<p>Intervention: Fast Track PATHS® Curriculum (based on published version by Kusche & Greenberg, 1995)</p> <p>Duration: 3 school years (Sept-May)</p> <p>Total PATHS lessons: First Grade: 57 Second Grade: 46 Third Grade: 48</p> <p>Lesson Duration: 20-30 mins</p> <p>Frequency: 2-3 times per week.</p> <p>Interventionist: Class Teacher.</p> <p>Control Condition: Not Reported.</p>	<p>Teacher Ratings: <i>Teacher Observation of Classroom Adaptation-Revised</i> (TOCA-R; Werthamer-Larsson, Kellam, & Wheeler, 1991).</p> <p><i>Social Health Profile</i> which assessed nine items describing prosocial behaviour and emotion regulation added to create a total score for social competence (SHP; CPPRG, 1998).</p> <p>Peer Nominations: Children nominated peers who fit behavioural descriptions and scores were standardised by classroom.</p> <p>Data Collection of Teacher Ratings: Term 1 & 2 of 1st Grade Term 2 of 2nd Grade Term 2 of 3rd Grade</p> <p>Data collection of Peer Nominations: No pre-test measure but collected in term 2 of 1st, 2nd, & 3rd Grade.</p>	<p>Findings: PATHS® intervention resulted in more positive outcomes for teacher rated social competence with small effect size (Cohen’s $d = 0.34$, $p < 0.001$)</p>

<p>Author: Novak et al. (2017)</p>	<p>Design: Cluster Randomised Controlled Trial.</p>	<p>Country: Croatia.</p>	<p>Sample size: 568</p>	<p>Intervention: PATHS® Curriculum translated into Croatian by translator.</p>	<p>Teacher Ratings: <i>Social Competence Scale</i> (http://www.fasttrackproject.org/techrept/s/sct/) provided ratings of prosocial behaviour; emotion regulation.</p>	<p>Findings:</p>
<p>Title: PATHS in Croatia: A school-based randomised-controlled trial of a social and emotional learning curriculum.</p>	<p>Groups: Random Allocation.</p>	<p>Setting: Mainstream Schools in Zagreb, Rijeka and Istria.</p>	<p>Age: 7 years old at the beginning of the study in the middle of First Grade and near the end of Second Grade when the study finished.</p>	<p>Duration: Last half of First Grade and first half of Second Grade (approximately one academic year).</p>	<p><i>School Readiness Questionnaire</i> (Bierman, Domitrovich et al., 2008) provided teacher ratings of learning behaviour.</p>	<p>PATHS® Intervention group did not show significant effects on improving children’s pro-social behaviour (Cohen’s $d = 0.16$, $p > 0.05$).</p>
	<p>Intervention: N = 280</p>		<p>Gender: 47% female.</p>	<p>Total PATHS lessons: 63</p>	<p><i>ADHD Rating Scale</i> (DuPaul, 1991) provided teacher ratings of inattention and Hyperactivity.</p>	
	<p>Control: N = 288</p>			<p>Lesson Duration: Not Reported.</p>	<p><i>TOCA-R</i> (Werthamer-Larsson et al., 1991) provided teacher ratings of oppositional behaviour and physical aggression.</p>	<p>PATHS® intervention group rated by teachers as showing a marginally greater improvement in emotional regulation compared to control group (Cohen’s $d = 0.18$, $p < 0.10$).</p>
				<p>Frequency: 2 per week.</p>	<p><i>Strengths and Difficulties Questionnaire (SDQ)</i> (Goodman, 1997; Goodman et al., 2010) provided teacher ratings on peer problems.</p>	
				<p>Interventionist: Class Teacher.</p>	<p>Data Collection: teachers completed pre-intervention ratings of child behaviour in Winter of First Grade. In Spring term of Second Grade, teachers completed the same post-intervention battery.</p>	
				<p>Control Condition: Usual Practice.</p>		

<p>Author: Curtis & Norgate (2007).</p>	<p>Design: Cohort Study.</p>	<p>Country: UK</p>	<p>Sample size: 287</p>	<p>Intervention: PATHS® Curriculum.</p>	<p>Teacher Ratings: <i>Strengths and Difficulties Questionnaire (SDQ)</i> (Goodman, 1997; Goodman et al., 2010) provided teacher ratings on emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems; prosocial behaviour.</p>	<p>Findings: Children in the PATHS® schools showed significant improvements on measures of prosocial behaviour on the SDQ (Cohen's $d = 0.91$, $p < 0.001$)</p>
<p>Title: An Evaluation of the Promoting Alternative Thinking Strategies Curriculum at Key Stage 1.</p>	<p>Groups: Non-Random Allocation. Control group due to receive intervention the following year.</p>	<p>Setting: Mainstream Primary School</p>	<p>Age: 5-7 years old. (Key Stage 1)</p>	<p>Duration: One Academic Year.</p>	<p>Teacher Interviews conducted and analysed qualitatively.</p>	<p>N.B. scores on the SDQ for intervention and control group were significantly different prior to intervention so caution needed when interpreting post-intervention results.</p>
	<p>Intervention: 5 schools N=114</p>		<p>Gender: Not Reported.</p>	<p>Total PATHS lessons: Not Reported.</p>	<p>Data Collection: SDQ completed at the beginning (pre-test) and end (post-test) of an academic year. SDQ completed by same teacher at both timepoints.</p>	
	<p>Control: 3 schools N=173</p>			<p>Lesson Duration: Not Reported.</p>	<p>Frequency: Not Reported.</p>	
				<p>Interventionist: Class Teacher.</p>	<p>Control Condition: Usual Practice.</p>	

<p>Author: Little et al. (2012)</p> <p>Title: The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK.</p>	<p>Design: Cluster Randomised Controlled Trial</p> <p>Groups: Random allocation of schools conducted independently of study authors.</p> <p>Intervention: 29 schools</p> <p>Control: 27 schools</p>	<p>Country: UK</p> <p>Setting: Mainstream Primary Schools in Birmingham, UK.</p>	<p>Sample size: 5,397 at baseline. 4,006 cases with all three sets of data.</p> <p>Age: Children aged 4-6 years (Reception, Year 1 and Year 2)</p> <p>Gender: Not Reported.</p>	<p>Intervention: PATHS® Curriculum.</p> <p>Duration: 2 academic years.</p> <p>Total PATHS lessons: Not Reported.</p> <p>Lesson Duration: 60 minutes</p> <p>Frequency: Once per week.</p> <p>Interventionist: Class Teacher.</p> <p>Control Condition: Waitlist control condition in which children received services as usual. In some schools this involved Social and Emotional Aspects of Learning (SEAL) programme.</p>	<p>Teacher Ratings: <i>Strengths and Difficulties Questionnaire (SDQ)</i> (Goodman, 1997; Goodman et al., 2010) provided teacher ratings on emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems; prosocial behaviour.</p> <p><i>PATHS Teacher Rating Survey (PTRS)</i> provided a composite of seven scales including the Child Behaviour Questionnaire. This assessed a range of behaviours such as emotion regulation, pro-social behaviour and social competence.</p> <p>Data Collection: Baseline in Sept 2009 First Follow up in June 2010 Second Follow up in June 2011</p>	<p>Findings: No significant effects of the PATHS® intervention on prosocial behaviour, emotional regulation or social-emotional competence were observed at the end of the second year of the intervention.</p>
--	---	--	--	--	---	--

<p>Author: Humphrey et al. (2018)</p>	<p>Design: Cluster Randomised Controlled Trial</p>	<p>Country: UK</p>	<p>Sample size: N=5,218</p>	<p>Intervention: PATHS® Curriculum.</p>	<p>Self-report Ratings: <i>Social Skills Improvement System</i> (Gresham & Elliot, 2008) provided a measure of the child's social skills.</p>	<p>Findings:</p>
<p>Title: The PATHS Curriculum for promoting social and emotional well-being among children aged 7-9 years: A cluster RCT.</p>	<p>Groups: Random allocation of independent of study authors.</p>	<p>Setting: Mainstream Primary Schools in Greater Manchester, UK.</p>	<p>Age: Children aged 7-9 years (Years 3-5 in UK schools)</p>	<p>Duration: 3 academic years.</p>	<p>Teacher Ratings: <i>Strengths and Difficulties Questionnaire (SDQ)</i> (Goodman, 1997; Goodman et al., 2010) provided teacher ratings on emotional symptoms; conduct problems; hyperactivity/inattention; peer relationship problems; prosocial behaviour.</p>	<p>No significant effects were found as a result of the PATHS® intervention on teacher reported pro-social behaviour (Cohen's $d = 0.07$, $p > 0.05$).</p>
	<p>Intervention: 23 schools (N=2294)</p>		<p>Gender: Intervention n = 49.9% Male.</p>	<p>Total PATHS lessons: Year 3 = 43 Year 4 = 57 Year 5 = 55</p>	<p>Data Collection: Baseline data collected May-July 2012 (T1) Interim (T2) Post-Intervention data collected May-July 2014 (T3)</p>	
	<p>Control: 22 schools (N=2106)</p>		<p>Control = 53.0% Male</p>	<p>Lesson Duration: 30-40 minutes</p>		
				<p>Frequency: Twice per week.</p>		
				<p>Interventionist: Class Teacher.</p>		
				<p>Control Condition: Schools in 'usual provision' had Social and Emotional Aspects of Learning (SEAL) programme and related interventions.</p>		

<p>Author: Malti et al. (2011)</p>	<p>Design: Cluster Randomised Controlled Trial</p>	<p>Country: Switzerland</p>	<p>Sample size: N=1,675 in a total of 56 schools.</p>	<p>Intervention of interest: PATHS® Curriculum culturally adapted to Swiss school system.</p>	<p>Teacher, parent and child ratings: <i>Social behaviour Questionnaire (Tremblay et al., 1991)</i> was completed by teachers, parents and children completed a self-report to give measures of social competence</p>	<p>Findings: PATHS® intervention resulted in no significant improvement in the intervention group's teacher rated pro-social behaviour (Cohen's $d = 0.14$, $p > 0.05$).</p>
<p>Title: The Effectiveness of Two Universal Preventative Interventions in Reducing Children's Externalising Behaviour: A Cluster Randomised Controlled Trial.</p>	<p>Groups: Random allocation using computer randomisation.</p>	<p>Setting: Mainstream Primary Schools in the city of Zurich.</p>	<p>Age: 7-11 years of age (1st year of Elementary school in Switzerland at start of study)</p>	<p>Duration: 1 School Year.</p>	<p>Data Collection: Data collection took place at annual intervals. T1: Allocation to treatment conditions T2: Baseline in 2nd Grade (2005/6) T3: Post Intervention 3rd Grade (2006/7) T4: Follow Up 5th Grade 2 years later (2008/9)</p>	
	<p>Intervention PATHS only 14 schools N=360</p>		<p>Gender: 48% Female</p>	<p>Total PATHS lessons: 46</p>		
	<p>Control: 14 schools N=356</p>			<p>Lesson Duration: 67 minutes</p>		
	<p>Triple P: 14 schools N=339</p>			<p>Frequency: 2.4 sessions per week</p>		
	<p>PATHS + Triple P 14 schools N=306</p>			<p>Interventionist: Class Teacher.</p>		
				<p>Control Condition: Not reported.</p>		

<p>Author: Seifer et al. (2004)</p>	<p>Design: Cohort Study.</p>	<p>Country: USA</p>	<p>Sample size: N=150</p>	<p>Intervention of interest: PATHS® Curriculum</p>	<p>Sociometric status: nominations for peer social status and social behaviour were collected. Principal Components Analysis (PCA) conducted limiting the number of factors to 2 – <i>Positive Peer Nominations</i> and <i>Negative Peer Nominations</i>.</p>	<p>Findings:</p>
<p>Title: Implementation of the PATHS Curriculum in an Urban Elementary School.</p>	<p>Groups: No details of randomisation procedure.</p> <p>Intervention: N=62 children.</p> <p>Control: N = 88 children.</p>	<p>Setting: Magnet Elementary school in inner-city province.</p>	<p>Age: 6-8 years of age (1st and 2nd Grade in US Schools)</p> <p>Gender: Not Reported.</p>	<p>Duration: 2 academic years</p> <p>Total PATHS lessons: Not reported.</p> <p>Lesson Duration: 30-45 minutes</p> <p>Frequency: 3 sessions per week</p> <p>Interventionist: Class Teacher.</p> <p>Control Condition: Treatment as usual</p>	<p>Emotion Understanding: <i>Kusche Affective Interview (KAI-R)</i> (Kusche, Greenberg, & Beilke, 1988) used to assess spontaneous emotion naming skills. <i>Kusche Emotional Inventory</i> (Greenberg & Kusche, 1998) was adapted to create a measure of emotion recognition.</p> <p>Social Status self-report: Children interviewed about their perceived status within the school environment. Included items such as ‘perceived meanness’, ‘perceived rejection’ and ‘negative feelings towards school’.</p>	<p>The PATHS® Intervention group demonstrated higher social-emotional competence compared to the control group Cohen’s $d = 0.38, p < 0.05$)</p>
<p>Child Depression Symptoms: <i>Childhood Depression Inventory</i> (CDI; Kovacs, 1992) was used to assess depressive symptoms.</p> <p>Global Social Competence: <i>Psychological Impairment Rating Scale</i> (PIRS; Gouley, Eguia, Seifer, Dickstein,</p>						

& Sameroff, 2003) was used following a 30 minute interview with the child to provide an overall rating of the child's social competence.

Social-Emotional Competence

Composite: composite measure computed by summing 9 standard scores detailed above.

Data Collection:

Baseline assessments not available for all cohorts so only post-test information was considered.

Compared Control group outcomes measures from 2001 with the outcome measures of the PATHS intervention group from 2002 (who had completed the intervention the previous academic year in 2001).

Appendix C: Rationale for Weight of Evidence A (WoE A)

Weight of Evidence A (WoE A) considered the methodological quality of the research. An adapted version of Kratochwill’s (2003) ‘group-based design coding protocol’ was used. It considered key methodological features including measurement, comparison groups, identifiable components, implementation fidelity and follow up assessment. The protocol was adapted for the current review and components of the protocol that were not considered relevant to the included studies were removed. Table 3 demonstrates the amendments made to the existing protocol.

Each study was assessed using the five features listed above and was given a score from ranging from 0-3 based on the criteria that they fulfilled within each category. Details as to the criteria for each methodological feature can be found in the Kratochwill coding manual. Each score for the five methodological features were then averaged to provide an overall WoE A rating.

Table 3: Modifications to Kratochwill (2003) coding protocol and rationale

Modified Section	Rationale
I. General Characteristics A – General study characteristics (removed)	This section provides a description of the study characteristics. As study characteristics are covered elsewhere in the review this was removed from the coding protocol.
II. Key features for Coding Studies and Rating Level of Evidence/Support A – Research Methodology (removed)	This section covered the methodological relevance of research which was addressed in WoE B and therefore removed from the coding protocol.
II. D – Primary/Secondary Outcomes Are Statistically Significant (removed)	The current review considered this in significant detail and so this feature was removed from the coding protocol.

II. E – Cultural Significance (removed)	The relevance of the culture in which the research was conducted was considered in WoE C and was therefore removed from the current coding protocol.
II. F – Educational/Clinical Significance (removed)	The current review considered this separately and so it was removed from the coding protocol.
II. G – External Validity Indicators (removed)	This section of the coding protocol was removed as the intervention investigated was a universal intervention and therefore inclusion/exclusion criteria did not apply to the samples involved in the study.
II. K – Replication (removed)	This section was removed from the current review as it was not an essential characteristic of the studies included.
II. L – Site of Implementation (removed)	This section was removed as all studies were conducted in school settings as specified in the inclusion/exclusion criteria.

Appendix D: WoE A Coding Protocol

Coding Protocol: Group-Based Design

Adapted from Kratochwill, T. R. (2003). Task Force on Evidence Based Interventions in School Psychology. American Psychological Association.

Domain:

- School- and community-based intervention programs for social and behavioral problems
- Academic intervention programs Family and parent intervention programs
- School-wide and classroom-based programs
- Comprehensive and coordinated school health services

Name of Coder(s): AH

Date: 29/12/2020

Full Study Reference in APA format: Bierman, K. L., Coie, J. D., Dodge, K. A., Greenberg, M. T., Lochman, J. E., McMahon, R. J., & Pinderhughes, E. (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.

Intervention Name (description from study): PATHS curriculum

Type of Publication: (Check one)

- Book/Monograph
- Journal article
- Book chapter
- Other (specify):

I. General Characteristics

B. General Design Characteristics

B1. Random assignment designs (if random assignment design, select one of the following)

- B1.1 Completely randomized design
- B1.2 Randomized block design (between-subjects variation)
- B1.3 Randomized block design (within-subjects variation)

B1.4 Randomized hierarchical design

B2. Nonrandomized designs (if nonrandom assignment design, select one of the following)

B2.1 Nonrandomized design

B2.2 Nonrandomized block design (between-participants variation)

B2.3 Nonrandomized block design (within-participants variation)

B2.4 Nonrandomized hierarchical design

B2.5 Optional coding of Quasi-experimental designs (see Appendix C)

B3. Overall confidence of judgment on how participants were assigned (select one of the following)

B3.1 Very low (little basis)

B3.2 Low (guess)

B3.3 Moderate (weak inference)

B3.4 High (strong inference)

B3.5 Very high (explicitly stated)

B3.6 N/A

B3.7 Unknown/unable to code

C. Data Analysis (answer B1 through B6)

C1. Appropriate unit of analysis yes No

C2. Familywise error rate controlled yes No N/A

C3. Sufficiently large *N* yes No

Statistical Test: multi-level logistic regression

α level: 0.05

ES: 0.20

N required: 988

C4. Total size of sample (start of the study): 2,937

C5. Intervention group sample size: 190 classes, *N* not specified

C6. Control group sample size: 180 classes, *N* not specified

D. Type of Program (select one)

- D1. Universal prevention program
- D2. Selective prevention program
- D3. Targeted prevention program
- D4. Intervention/Treatment
- D5. Unknown

E. Stage of the Program (select one)

- E1. Model/demonstration programs
- E2. Early stage programs
- E3. Established/institutionalized programs
- E4. Unknown

F. Concurrent or Historical Intervention Exposure (select one)

- F1. Current exposure – Fast Track
- F2. Prior exposure
- F3. Unknown

II. Key Features for Coding Studies and Rating Level of Evidence/Support

(3=Strong Evidence 2=Promising Evidence 1=Weak Evidence 0=No Evidence)

B. Measurement (answer B1 through B6)

B1. Use of outcome measures that produce reliable scores for the majority of primary outcomes. The table for Primary/Secondary Outcomes Statistically Significant allows for listing separate outcomes and will facilitate decision making regarding measurement (select one of the following)

- B1.1 Yes
- B1.2 No
- B1.3 Unknown/unable to code

B2. Multi-method (select one of the following)

- B2.1 Yes

- B2.2 No
- B2.3 N/A
- B2.4 Unknown/unable to code

B3. Multi-source (select one of the following)

- B3.1 Yes
- B3.2 No
- B3.3 N/A
- B3.4 Unknown/unable to code

B5. Validity of measures reported (select one of the following)

- B5.1 Yes validated with specific target group
- B5.2 In part, validated for general population only
- B5.3 No
- B5.4 Unknown/unable to code

B. Overall Rating for Measurement (select 0, 1, 2, or 3):

- 3 2 1 0

C. Comparison Group

C1. Type of Comparison Group (select one of the following)

- C1.1 Typical contact
- C1.2 Typical contact (other) specify:
- C1.3 Attention placebo
- C1.4 Intervention elements placebo
- C1.5 Alternative intervention
- C1.6 Pharmacotherapy
- C1.7 No intervention
- C1.8 Wait list/delayed intervention
- C1.9 Minimal contact
- C1.10 Unable to identify comparison group

C2. Overall confidence rating in judgment of type of comparison group (select one of the following)

- C2.1 Very low (little basis)
- C2.2 Low (guess)
- C2.3 Moderate (weak inference)
- C2.4 High (strong inference)
- C2.5 Very high (explicitly stated)
- C2.6 Unknown/Unable to code

C3. Counterbalancing of Change Agents (answer C3.1 to C3.3)

- C3.1 By change agent
- C3.2 Statistical
- C3.3. Other

C4. Group Equivalence Established (select one of the following)

- C4.1 Random assignment
- C4.2 Post hoc matched set
- C4.3 Statistical matching
- C4.4 Post hoc test for group equivalence

C5. Equivalent Mortality (answer C5.1 through C5.3)

- C5.1 Low Attrition (less than 20% for Post)
- C5.2 Low Attrition (less than 30% for follow-up)
- C5.3 Intent to intervene analysis carried out

Findings _____

C. OVERALL Rating for Comparison Group (select 0, 1, 2, or 3):

- 3 2 1 0

H. Durability/Generalization of Intervention and Outcomes

H1. Follow-up assessment

H1.1 Timing of follow up assessment: yes no

Specify _____

H1.2. Number of participants included in the follow up assessment: N/A

H1.3, Consistency of assessment method used: N/A

H1.4. Follow-up addresses institutionalization and/or sustainability of intervention efforts: N/A

H. Overall Rating for Follow-up Assessment (select 0, 1, 2, or 3):

3 2 1 0

I. Identifiable Intervention Components (answer I1 through I7)

I1. Overall Rating for Identifiable Components: 3 2 1 0

I2. Design allows for analysis of identifiable components (select one)

yes no

I3. Total number of components: N/A

I4. Number of components linked to primary outcomes: N/A

I5. Clear documentation of essential components (select one) yes
no

I6. Procedures for adapting the intervention are described in detail (select one)

yes no

I7. Contextual features of the intervention are documented (select one)

yes no

I. OVERALL Rating of Identifiable Intervention Components (select 0, 1, 2, or 3):

3 2 1 0

J. Implementation Fidelity

J1. Evidence of Acceptable Adherence (answer J1.1 through J1.3)

J1.1 Ongoing supervision/consultation

J1.2 Coding intervention sessions/lessons or procedures

J1.3 Audio/video tape implementation (select J1.3.1 or J1.3.2):

J1.3.1 Entire intervention

J1.3.2 Part of intervention

J2. Manualization (select all that apply)

J2.1 Written material involving a detailed account of the exact procedures and the sequence in which they are to be used

J2.2 Formal training session that includes a detailed account of the exact procedures and the sequence in which they are to be used

J2.3 Written material involving an overview of broad principles and a description of the intervention phases

J2.4 Formal or informal training session involving an overview of broad principles and a description of the intervention phases

J3. Adaptation procedures are specified (select one)

yes

no

unknown

Rating for Implementation Fidelity (select 0, 1, 2, or 3):

3 2 1 0

III. Other Descriptive or supplemental Criteria to Consider

G. External Validity Indicators

G1. Sampling Procedures

G1.1 Sampling procedures described in detail yes no

Inclusion/exclusion criteria specified yes no

Inclusion/exclusion criteria similar to school practice yes no

Specified criteria related to concern yes no

G2. Participant Characteristics Specified for Treatment and Control Group (modified):

Age/School Year

Gender

SEN diagnostic label

Ethnicity

Home Language

Socio-economic background

- Levels of general cognitive abilities
- Levels of social difficulties
- Levels of emotional difficulties

G3. Details are provided regarding variables that:

G3.1 Have differential relevance for intended outcomes yes
no

Specify: High risk pupils based on social emotional difficulties and location of intervention.

G3.2 Have relevance to inclusion criteria yes no

H2. Durability/Generalization over time

H2.1.1 Evidence is provided regarding the sustainability of outcomes after intervention is terminated yes no

H2.1.2 Procedures for maintaining outcomes are specified yes no

H3. Generalization across settings

H3.1 Evidence is provided regarding the extent to which outcomes are manifested in contexts that are different from the intervention context:
yes no

H3.2 Documentation of efforts to ensure application of intervention to other settings.
yes no

H3.3 Impact on implementers or context is sustained
yes no

H4. Generalization across persons

Evidence is provided regarding the degree to which outcomes are manifested with participants who are different than the original group of participants for with the intervention was evaluated

- yes
- no

J4.5 Length of Intervention (select J4.5.1 or J4.5.2)

- Unknown/insufficient information provided
- Information provided (if information is provided, specify one of the following:)
- weeks _____
- months _____

years 3 academic years

other _____

J4.6 Intensity/dosage of Intervention (select J4.6.1 or J4.6.2)

Unknown/insufficient information provided

Information provided (if information is provided, specify both of the following:)

length of intervention session 20-30 minutes

frequency of intervention session 2-3 times per week

J4.8 Program Implementer (select all that apply)

Research Staff

School Specialty Staff

Teachers

Educational Assistants

Parents

College Students

Peers

Other

Unknown/insufficient information provided

J4.11 Training and Support Resources (select all that apply)

Simple orientation given to change agents

Training workshops conducted

of Workshops provided: 2

Average length of training: Full Day

Who conducted training (select all that apply)

Project Director

Graduate/project assistants

Other (please specify): Certified PATHS trainer

Unknown

Ongoing technical support

Program materials obtained

Special Facilities

Other (specify):

Summary of Evidence for Group-Based Design Studies

Indicator	Overall Evidence Rating 0-3 Or NNR = No Numerical Rating	Description of Evidence Strong Promising Weak No/Limited Evidence Or Descriptive Rating
<i>Key areas of judgement for Weight of Evidence A</i>		
Measurement	2	Promising
Comparison Group	1	Weak
Identifiable Component	0	No Evidence
Implementation Fidelity	3	Strong
Follow up assessment conducted	0	No Evidence

Average Quality of Evidence across the Key Judgement Areas		
$\frac{\Sigma \text{ of } X}{N}$	$\frac{2+1+0+3+0}{5}$	$= 1.2$
<p>X= Individual quality of evidence for each judgement area</p>		
<p>N= Number of judgement areas</p>		
<p>Overall Rating for Weight of Evidence A: 1.2 (Low)</p>		

Appendix E: Summary Table of WoE A

Table 4: Overall WoE A scores for studies included in the review

Reference	Measurement	Comparison Group	Identifiable Component	Implementation Fidelity	Follow Up Assessment	Overall WoE A*
Bierman et al. (2010)	2	1	0	3	0	1.2
Novak et al. (2017)	2	2	0	2	0	1.2
Curtis & Norgate (2007)	2	2	0	2	0	1.2
Little et al. (2012)	2	2	0	1	0	1.4
Humphrey et al. (2018)	3	1	0	3	1	1.6
Seifer et al. (2004)	1	1	0	1	0	0.6
Malti et al. (2011)	3	2	0	3	0	2.0

*calculated using the average score by adding up the rating for each individual category and dividing the total by 5

Note 1: WoE ratings receive a rating of low if ≤ 1.4 , medium if between 1.5-2.4 and high if ≥ 2.5 .

Appendix F: Rationale and Criteria for Weight of Evidence B (WoE B)

WoE B considered the methodological relevance of research and its appropriateness to the review question based on the recommendations of Petticrew and Robert’s typologies of evidence (2003).

Table 5: WoE B Criteria

Criteria	1 – Low	2 – Medium	3 – High
Study design	Non-experimental design.	Quasi-experimental studies or cohort studies with no randomization to groups.	Randomised Controlled Trials.
Comparison Group	Single group design with no control group.	At least one ‘no intervention’ comparison group in which the control group is on a waitlist or is in a minimal contact condition.	A minimum of one ‘active’ control group in which the control group receives an alternative intervention or attention placebo.
Reporting of data collection	Outcome measures are only reported post intervention.	Outcome measures are provided for one group only at pre- and post-intervention.	Outcomes measures are reported pre- and post-intervention for intervention and control group.

Reliability and Validity of outcome measures	Outcome measures to test social-emotional competence are described but no details provided regarding reliability or validity of the measures.	Reliability or validity data is provided for outcomes used to assess social-emotional competence.	Reliability and validity data are provided for outcomes to assess social-emotional competence.
--	---	---	--

Appendix G: Summary Table of WoE B

Table 6: Overall WoE B scores for studies included in the review

Study	Study Design	Comparison Group	Reporting of data collection	Reliability and Validity of outcome measures	Overall WoE B *
Bierman et al. (2010)	3	2	3	2	2.5 (high)
Novak et al. (2017)	3	2	3	2	2.5 (high)
Curtis & Norgate (2007)	2	2	3	1	2.0 (medium)
Little et al. (2012)	3	3	3	1	2.5 (high)
Humphrey et al. (2018)	3	3	3	3	3.0 (high)
Seifer et al. (2004)	2	2	1	2	1.8 (medium)
Malti et al. (2011)	3	2	3	2	2.5 (high)

**calculated using the average score by adding up the rating for each individual category and dividing the total by 4.*

Note 1: WoE ratings receive a rating of low if ≤ 1.4 , medium if between 1.5-2.4 and high if ≥ 2.5 .

Appendix H: Rationale and Criteria for Weight of Evidence C (WoE C)

WoE C evaluated the topic relevance of each study included in the current review. Four relevant topic areas were evaluated including the intervention location, participant characteristics, implementation fidelity and the setting of the intervention.

Table 7: WoE C Criteria

Criteria	1 – Low	2 – Medium	3 - High	Rationale
Intervention Location	Interventions took place in schools that are not located in OECD countries.	Intervention took place in schools within OECD countries.	Intervention took place in UK primary schools.	The location of the intervention is relevant as the current setting is UK schools. Studies that took place in UK schools are therefore more relevant to Educational Psychology practice as the findings are likely to be more similar to the current context.
Participant Characteristics Table 8 provides details of participant characteristics that could be included in each study.	The study reports on 0-2 of the participant characteristics.	The study reports on 3-5 of the participant characteristics.	The study reports on 6-7 of the participant characteristics.	The more details that are provided regarding the participants in the study, the more likely it is that information about the generalisability of the intervention is available.

Implementation Fidelity	Self-report of fidelity with no additional support from independent source or no measures of implementation fidelity.	Self-report of fidelity – such as how often lessons were delivered or how closely the manual was followed and access to technical support.	Independent observation and feedback provided to teachers delivering the intervention with opportunities for ongoing support and top-up training.	The degree to which the school staff received support in delivering the intervention is likely to influence the quality with which it was delivered and is likely to impact the outcomes of the intervention.
Setting	Participants were selected from a non-educational sample.	Participants were selected from one mainstream primary school.	Participants were selected from multiple mainstream primary schools.	The study will have greater external validity if the intervention is implemented across multiple sites and the findings are replicated.

Table 8: Participant Characteristics detailed in each included study

Participant Characteristic	Bierman et al. (2010)	Novak et al. (2017)	Curtis & Norgate (2007)	Little et al. (2012)	Humphrey et al. (2018)	Seifer et al. (2004)	Malti et al. (2011)
Age or School Year	√	√	√	√	√	√	√
Gender		√			√		
Location of school setting (e.g. urban vs rural)	√			√		√	√
Ethnicity	√				√	√	
Socio-economic Status	√				√	√	√
Languages Spoken					√	√	
Special Educational Needs					√		√

Appendix I: Summary Table of WoE C

Table 9: Overall WoE C scores for studies included in the review

Study	Intervention Location	Participant Characteristics	Implementation Fidelity	Setting	Overall WoE C *
Bierman et al. (2010)	2	2	3	3	2.5 (high)
Novak et al. (2017)	1	1	3	3	2.0 (medium)
Curtis & Norgate (2007)	3	1	2	3	2.3 (medium)
Little et al. (2012)	3	1	2	3	2.3 (medium)
Humphrey et al. (2018)	3	3	3	3	3.0 (high)
Seifer et al. (2004)	2	2	1	2	1.8 (medium)
Malti et al. (2011)	2	2	3	3	2.5 (high)

*Calculated using the average score by adding up the rating for each individual category and dividing the total by 4.

Note 1: WoE ratings receive a rating of low if ≤ 1.4 , medium if between 1.5-2.4 and high if ≥ 2.5 .

Appendix J: Overall Weight of Evidence D

The overall ratings for WoE A, B and C were given equal weightings and were added together and divided by three to give an overall WoE D rating. WoE D provides a summary of the overall strength of the evidence in answering the current review question and qualitative descriptors can be found in Table 11.

Table 10: Overall WoE D scores for studies included in the review

Study	WoE A Methodological Quality	WoE B Methodological Relevance	WoE C Topic Relevance	WoE D Overall weight of evidence
Bierman et al. (2010)	1.2	2.5	2.5	2.1 Medium
Novak et al. (2017)	1.2	2.5	2.0	1.9 Medium
Curtis & Norgate (2007)	1.2	2.0	2.3	1.8 Medium
Little et al. (2012)	1.4	2.5	2.3	2.1 Medium
Humphrey et al. (2018)	1.6	3.0	3.0	2.5 High
Seifer et al. (2004)	0.6	1.8	1.8	1.4 Low
Malti et al. (2011)	2.0	2.5	2.5	2.3 Medium

Table 11: Qualitative descriptions of overall WoE D ratings

Overall WoE Rating	
≤ 1.4	Low
1.5-2.4	Medium
≥ 2.5	High