Bangor University

DOCTOR OF PHILOSOPHY

Exploring ways to increase the use of evidence-based methods in a cluster of schools in Wales

Pegram, Jane

Award date:
2023

Awarding institution:
Bangor University

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 10. Apr. 2024

Prifysgol Bangor

Bangor University

Exploring ways to increase the use of evidence-based methods in a cluster of schools in Wales

Jane Olwen Pegram

Thesis submitted to the School of Educational Sciences, Bangor University, in partial fulfilment for the degree of Doctor of Philosophy

September 2022

Declaration

‘I hereby declare that this thesis is the results of my own investigations, except where otherwise stated. All other sources are acknowledged by bibliographic references. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree unless, as agreed by the University, for approved dual awards.

I confirm that I am submitting this work with the agreement of my Supervisor(s).’

‘Yr wyf drwy hyn yn datgan mai canlyniad fy ymchwil fy hun yw’r thesis hwn, ac eithrio lle nodir yn wahanol. Caiff ffynonellau eraill eu cydnabod gan droednodiadau yn rhoi cyfeiriadau eglur. Nid yw sylwedd y gwaith hwn wedi cael ei dderbyn o’r blaen ar gyfer unrhyw radd, ac nid yw’n cael ei gyflwyno ar yr un pryd mewn ymgeisiaeth am unrhyw radd oni bai ei fod, fel y cytunwyd gan y Brifysgol, am gymwysterau deuol cymeradwy.

Rwy'n cadarnhau fy mod yn cyflwyno'r gwaith hwn gyda chytundeb fy Ngoruchwyliwr (Goruchwylwyr)

Acknowledgements

This PhD would not have been possible without the support and assistance of many people. First and foremost, I would like to thank my project supervisors Professor Carl Hughes, Dr Richard Watkins, and Dr Marguerite Hoerger for not only providing the opportunity for me to conduct such innovative and impactful research, but also for their academic support, mentorship, guidance, time and patience. Carl, thank-you for providing me with many opportunities to develop my skills, knowledge and confidence in research. Richard, thank you for nurturing the research partnership and your invaluable support and feedback throughout the whole process. Maggie, thank you for providing invaluable input and expertise.

Throughout the duration of this PhD, I was fortunate enough to work with and be supported by a cluster of wonderful head teachers, who not only used their scarce resources to part fund this PhD studentship, but also contributed their time and ideas to help steer the research – without you this thesis would not be possible. In particular, thank you to John Cannon for your time, support, co-operation and enthusiasm. Furthermore, I would like to thank all the teachers involved, who gave up their valuable time to contribute – without your cooperation, willingness and enthusiasm I would not have been able to conduct this research. And, not forgetting the spirited pupils that contributed – I am grateful for you all.

There are many other people that have supported and guided me through this PhD and to everyone who has played a part I am extremely grateful. A big thank you to GwE who part funded the studentship and supported the project throughout.

Finally, a big thank you to friends and colleagues that offered words of encouragement and helped me to keep pushing on.

Table of Contents

DECLARATION ..... **II**

ACKNOWLEDGEMENTS ..... **III**

SUMMARY **1**

CHAPTER 1: INTRODUCTION **3**

BENEFITS OF HIGH QUALITY EDUCATION 4
DEFINITIONS AND REVIEW OF AN EVIDENCE-INFORMED APPROACH 4
THE BARRIERS TO AN EVIDENCE-INFORMED APPROACH 7
THE FACILITATORS OF AN EVIDENCE-INFORMED APPROACH 9
ADOPTION 11
DIFFUSION OF INNOVATIONS THEORY 11
IMPLEMENTATION 12
WELSH EDUCATION SYSTEM AND CURRICULUM FOR WALES 14
SOCIAL AND EMOTIONAL APPROACHES IN SCHOOLS 17
GOOD BEHAVIOR GAME (GBG) 18
PROMOTING ALTERNATIVE THINKING STRATEGIES (PATHS) 19

OVERVIEW OF THESIS **21**

CHAPTER 2: METHODS AND DESIGN **26**

CHAPTER 3: ASSESSING THE RANGE AND EVIDENCE-BASE OF INTERVENTIONS IN A CLUSTER OF SCHOOLS. **33**

ABSTRACT **34**

METHOD **44**

Design and analysis *44*
Participants *45*
Data collection methods *45*
Procedure *46*
Screening of the interventions *47*
Definition of a school-based intervention *48*
Categorising the interventions *50*
RAPID SYSTEMATIC REVIEW OF THE EVIDENCE 51
Literature selection *52*
Screening and data extraction *53*
Quality assessment *54*
EVIDENCE SYNTHESIS AND PRESENTATION OF THE RESULTS 54
RESULTS 54
Cognition and Learning *55*
Accelerated Reader (Renaissance Learning) *66*
Read Write Inc. Fresh Start *68*
Friends for Life (FRIENDS) *69*
Picture Exchange Communication System (PECS) *74*
Intervention programmes and approaches with preliminary evidence of causal effect or evidence of ineffectiveness. *75*
Catch-Up Literacy *76*
Headsprout *76*
The Reading Educational Assistance Dogs (R.E.A.D) programme *78*
Toe by Toe *78*
Incredible Years, Dinosaur School *78*

Mindfulness in Schools Project, Paws b *79*
Time-out *80*
Student Assistance Program (SAP) *80*
Language for Thinking *80*
Social Stories *81*
A Narrative Intervention by Becky Shanks *81*
Lego® -Based Therapy *82*
RESULTS FROM THE FOLLOW-UP INVESTIGATION 82
Compatibility and impact *84*
Greater awareness of the evidence *85*
DISCUSSION 85
Directions for research *90*
Conclusions *90*

REFERENCES FROM THE RAPID SYSTEMATIC REVIEW 92

CHAPTER 4: THE FACTORS THAT INFORM SCHOOL LEADERS' CHOICE OF TEACHING PROGRAMMES 113

ABSTRACT 114

METHOD 120

ETHICS 120
PARTICIPANTS 120
INTERVIEW DESIGN AND PROCEDURE 120
ANALYSIS 121

RESULTS 123

THEME: EXPERIENCE AND ADVICE FROM EDUCATION PROFESSIONALS 123
Advice and experience from within school *123*
Advice and experience from other schools *124*
Advice from regional consortia and local authority *125*
THEME: COMPATIBILITY 126
Compatible with pupil needs *127*
Compatible with school needs *127*
THEME: SALES AND MARKETING 129
THEME: ACCESSING THE EVIDENCE-BASE 129

DISCUSSION 129

CHAPTER 5: EXPLORING THE FACTORS THAT AFFECT ADOPTION AND IMPLEMENTATION OF THE GOOD BEHAVIOR GAME, FOLLOWING TYPICAL IN-SERVICE TRAINING 136

METHOD 145

CONTEXT 145
GOOD BEHAVIOUR GAME TRAINING 145
GOOD BEHAVIOUR GAME 146
ETHICS 147
PARTICIPANTS 148
Outcome measures *149*
Observations *150*
Data collection procedure *151*
ANALYSIS 152

RESULTS 153

PRIOR CONDITIONS 155
Felt needs/problems *155*

Innovativeness *156*
Continuing professional development *156*
KNOWLEDGE PHASE 156
PERSUASION PHASE 156
Relative advantage *156*
Compatibility *157*
Observability *158*
Change agent *158*
DECISION PHASE 158
IMPLEMENTATION AND CONFIRMATION PHASES 158

DISCUSSION 159

LIMITATIONS 162
IMPLICATIONS 163

CHAPTER 6: A CULTURAL AND CONTEXTUAL ADAPTION OF THE PATHS PROGRAMME TO ENHANCE IMPLEMENTATION AND ACCEPTABILITY 165

ABSTRACT 166

METHOD 174

ETHICS 174
PATHS® (UK) 175
PROCESS OF ADAPTING PATHS 176

PRELIMINARY FEASIBILITY EVALUATION 177

INTERVENTION 177
TRIAL DESIGN AND PARTICIPANTS 178
PROCEDURE 180
MEASURES 180
OBSERVATIONS 182
FOCUS GROUPS 183
ANALYSIS 183

RESULTS 184

THEME 1: APPROACHES TO SEL 187
Social and Emotional Aspects of Learning (SEAL) *187*
Other approaches *188*
THEME 2: COMPATIBLE WITH BELIEFS, VALUES AND NEED 188
THEME 3: ACCEPTABLE AND SUITABLE 189
Resources and materials *189*
Structure and content *189*
Dosage and delivery time *190*
THEME 4: DISSATISFACTION 190
THEME 5: NEGATIVE IMPACT FROM COMPETING PRIORITIES 190
THEME 6: FURTHER ADAPTIONS 190
Use of additional resources *191*
Minor structural modifications *191*
THEME 7: INTENTIONS TO SUSTAIN USE 191
THEME 8: NEGATIVE IMPACT FROM COMPETING PROGRAMMES 192

DISCUSSION 192

Limitations *198*
Implications *200*

Conclusion *201*

CHAPTER 7: DISCUSSION OF THESIS **202**

DISCUSSION **203**

BROAD OVERVIEW 203
CHAPTER 3 204
CHAPTER 4 206
CHAPTER 5 209
CHAPTER 6 212

IMPLICATIONS **215**

STRENGTHS **217**

LIMITATIONS **218**

CONCLUSION **220**

List of Figures for Chapter 3

FIGURE 1. FLOW DIAGRAM DISPLAYING THE NUMBER OF INTERVENTIONS INCLUDED AND EXCLUDED FROM THE STUDY 49
FIGURE 2 FLOW DIAGRAM ILLUSTRATING STUDY SELECTION PROCESS (ADAPTED FROM PRISMA DIAGRAM—MOHER ET AL., 2009) 54

List of Figures for Chapter 4

FIGURE 1. THEMES DERIVED FROM THE DATA FOR QUESTION 1 (THINK OF A LITERACY OR NUMERACY PROGRAMME THAT YOU RECENTLY ADOPTED AND EXPLAIN WHY YOU CHOSE IT) 123
FIGURE 2. THEMES DERIVED FROM THE DATA FOR QUESTION 2 (THINK OF A WELL-BEING PROGRAMME THAT YOU RECENTLY ADOPTED AND EXPLAIN WHY YOU CHOSE IT) 123

List of Figures for Chapter 5

FIGURE 1. THEORETICAL FRAMEWORK BASED ON ROGERS (2003) INNOVATION-DECISION PROCESS 148

List of Tables

List of Tables for Chapter 3

TABLE 1 *EXAMPLES OF INTERVENTIONS EXCLUDED FROM THE STUDY, FOLLOWING SCREENING* 49
TABLE 2. *FREQUENCY OF INTERVENTIONS IN EACH SCHOOL ORGANISED BY SEN 'AREAS OF NEED'* 55
TABLE 3. *COGNITION AND LEARNING INTERVENTIONS USED ACROSS THE CLUSTER* 56
TABLE 4. *SOCIAL, EMOTIONAL & MENTAL HEALTH INTERVENTIONS USED ACROSS THE CLUSTER* 58
TABLE 5. *COMMUNICATION AND INTERACTION INTERVENTIONS USED ACROSS THE CLUSTER* 59
TABLE 6. *SUMMARY OF THE CAUSAL EVIDENCE FOR INTERVENTIONS USED ACROSS THE CLUSTER* 60
TABLE 7. *SUMMARY OF THE CAUSAL* EVIDENCE *FOR THE COGNITION AND LEARNING INTERVENTIONS* 62
TABLE 8. *SUMMARY OF THE CAUSAL EVIDENCE FOR THE SOCIAL, EMOTIONAL AND MENTAL HEALTH INTERVENTIONS* 63
TABLE 9. *SUMMARY OF THE CAUSAL EVIDENCE FOR THE COMMUNICATION AND INTERACTION INTERVENTIONS* 65
TABLE 10. *A SUMMARY OF THE EVIDENCE FOR SAFMEDS* 66
TABLE 11. *A SUMMARY OF THE EVIDENCE FOR ACCELERATED READER* 68
TABLE 12. *A SUMMARY OF THE EVIDENCE FOR READ WRITE INC. FRESH START* 68

TABLE 13. *A SUMMARY OF THE EVIDENCE FOR FRIENDS FOR LIFE (FRIENDS)* 73
TABLE 14. *A SUMMARY OF THE EVIDENCE FOR THE PICTURE EXCHANGE COMMUNICATION SYSTEM (PECS)* 75
TABLE 15. *A SUMMARY OF THE EVIDENCE FOR CATCH-UP LITERACY* 76
TABLE 16. *INTERVENTIONS IN USE IN PRIMARY SCHOOL X AT FOLLOW-UP* 83

List of Tables for Chapter 5

TABLE 1. *AN OUTLINE OF THE PROCEDURE FOR THE GOOD BEHAVIOR GAME* 147
TABLE 2. *CHARACTERISTICS OF THE TEACHERS INCLUDED IN THE STUDY* 148
TABLE 3. *DISPLAYS THE BEHAVIOUR MANAGEMENT TECHNIQUES OR STRATEGIES CURRENTLY USED BY THE TEACHERS* 149
TABLE 4. *FREQUENCY OF TEACHERS THAT PARTICIPATED IN THE DATA COLLECTION METHODS* 152
TABLE 5. *ADOPTION DECISIONS AND TOTAL SCORES FOR TEACHERS FOR QUESTION 11 AND 12 FROM THE TRAINING EVALUATION FORM* 153
TABLE 6. *MEAN SCORES FOR THE QUESTIONS FROM THE TRAINING EVALUATION QUESTIONNAIRE* 154
TABLE 7. *TOTAL ACCEPTABILITY SCORES ON THE IRP-15 FOR THE TEACHERS THAT ADOPTED AND ACTIVELY REJECTED THE GBG* 154
TABLE 8. *RESULTS FROM THE OBSERVATION CONDUCTED WITH ONE TEACHER* 155

List of Tables for Chapter 6

TABLE 1. *AN OUTLINE OF THE SYSTEMATIC PROCESS OF ADAPTION* 176
TABLE 2. *CHARACTERISTICS OF THE TEACHERS AND FREQUENCY IN EACH GROUP* 179
TABLE 3. *CHARACTERISTICS OF THE PUPILS AND FREQUENCY IN EACH GROUP* 179
TABLE 4. *DISPLAYS THE AVERAGE IMPLEMENTATION SCORES FOR DOSAGE, FIDELITY, QUALITY AND RESPONSIVENESS, ACROSS THE INTENDED 12 WEEKS OF LESSONS* 185
TABLE 5. *DISPLAYS THE DOSAGE, FIDELITY, QUALITY AND RESPONSIVENESS DATA DERIVED FROM THE OBSERVATIONS* 185
TABLE 6. *RESULTS FROM THE TEACHER-REPORT PUPIL EMOTION REGULATION CHECKLIST* 186
TABLE 7. *RESULTS FROM THE TEACHER-REPORT PUPIL SOCIAL COMPETENCE SCALE (SCS)* 187
TABLE 8. *RESULTS FROM THE TEACHER-REPORT PUPIL STRENGTHS AND DIFFICULTIES QUESTIONNAIRE (SDQ)* 187

Summary

To improve the quality and equity of the education system in Wales, there has been a renewed focus on creating a more evidence-informed practice. Within a self-improving system, teachers and school leaders are now largely responsible for improving the quality of provision by identifying and developing best practice through teacher enquiry and the use of external research evidence. However, the ways in which evidence-into-use is to be achieved is yet to be explicitly defined. The empirical research presented in Chapters 3, 4, 5 and 6 were conducted with the staff and pupils from a well-established cluster of schools in Wales. The cluster had entered a collaborative research partnership with GwE (regional consortia) and CIEREI, Bangor University, to commission close-to-practice research that aligns with the values and priorities of the organisations involved. The primary aim for the partnership was to increase the use of evidence-based methods in the cluster and move them towards a more evidence-informed practice. A priority outlined by the senior leaders in the schools was to improve their understanding of the evidence base for the interventions already in use (Chapter 3). The main research aims set out by GwE, and Bangor University were to: explore ways to increase the use of evidence-based methods in schools (which included understanding what facilitates senior leaders adoption-decisions – Chapter 4); exploring the effectiveness of educator-friendly reviews of evidence on existing provision (Chapter 3); evaluating the extent that teachers implement evidence-based programmes following typical in-service training (Chapter 5); and, finally, to describe the process of contextually adapting an evidence-based programme to enhance compatibility, and to evaluate if this leads to increased adoption and high levels of implementation (Chapter 6). Our results suggest that schools adopt a large number of interventions, which impact school budgets, teachers' time and capacity to adopt and use other evidence-based methods, and the extent that provision is implemented. We discovered that research evidence is infrequently used to inform school provision, and that many programmes and teaching methods lack a robust evidence-base. When choosing provision, we found that school leaders more often seek information from other educators and the local school improvement advisor, and, are more likely to adopt provision of it is perceived as compatibility with context. When exploring ways to increase the use of evidence-based methods, we found that educator-friendly summaries of evidence have little impact on the provision that schools use, and typical in-service training is not effective. However, the adaption of an evidence-based programme to make it more contextually compatible, does lead to adoption, high levels of implementation and intentions to sustain use. The evidence derived from the empirical chapters has provided important new information about the range of interventions in use in a cluster of school in Wales; contributed to the literature on determining how best to enhance the use of evidence-based methods in schools; and improved the quality of knowledge and decision-making by senior leaders, moving the cluster towards more evidence-informed practice. The limitations of the research and implications for the cluster, policy and research are discussed in the concluding chapter of this thesis.

Chapter 1: Introduction

Benefits of high quality education

Literacy and numeracy are universally acknowledged to be one of the essential foundations of education, and the evidence suggests that increased skills have a large positive effect on a person's academic attainment, future employment, earnings, and perceived quality of life. It is also acknowledged that increased literacy and numeracy skills impact on the collective skills of a workforce, employment rates, and economic growth (Campbell 1980; Currie & Thomas, 1999; Green & Riddell, 2001; Lleras, 2008; Machin et al., 2001). Hanushek and Woessmann (2015) provide evidence to suggest that it is the cognitive skills acquired through a school education that have powerful effects on an individual's outcomes and the economy. According to Hanushek and Woessmann (2020) it is not just the time spent in school that matters but rather the quality of knowledge that is acquired whilst there. However, research also suggests that cognitive ability alone is less predictive of future success (Levin, 2012). Levin proposes that cognitive, and non-cognitive abilities such as motivation, self-control, and pro-social behaviours, which can be taught through social and emotional curricula, are essential for educational attainment, and future life success. For example, children and young people that are socially and emotionally competent are more able to focus, maintain attention on tasks, and follow instruction in the classroom. And, they are more likely to develop the necessary life skills required to plan, set goals, make decisions, co-operate, and cope with adversity (Greenberg et al., 2017; Jones, Greenberg & Crowley, 2015; Panayiotou, Humphrey & Wigelsworth, 2019).

Definitions and review of an evidence-informed approach

To improve the quality and equity of the education system, education policymakers in jurisdictions across the United Kingdom have renewed their focus on an 'evidence-based' or 'evidence-informed' approach to teaching and learning, which is reflected in a number of significant reports (Donaldson, 2015; DfE, 2010, 2016; Institute for Effective Education, 2019; OECD, 2014; 2017; Welsh Government, 2014). For example, the Welsh Government's

National Strategy for Educational Research and Enquiry (NSERE): vision document (Welsh Government, 2021) represents a commitment to ensuring that education policy and practice in Wales is informed by the best available research evidence and findings from teacher enquiry. Within a self-improving system (as it is known in England) or a professional learning organisation (as it is known in Wales), teachers and school leaders are largely responsible for improving the quality of provision by identifying and developing best practice through teacher enquiry and the use of external research evidence (Gorard, 2020; Greany, 2015; OECD, 2018; Slavin, 2020; Welsh Government, 2017). The approach to evidence-use adopted within these education systems is an ‘evidence-informed’ approach (Coldwell et al., 2017; Welsh Government, 2021). The term evidence-informed was coined to better represent the multiple sources of knowledge and information that are required when decision making (Sharples, 2013) that arguably, an evidence-based approach neglects to achieve (Biesta 2007). An evidence-based approach to decision making considers only the external research evidence. An evidence-informed approach integrates the ‘best’ available evidence from worldwide research, the professional judgement and experience of educators and system level and classroom data, to improve practice (Gorard, See, & Siddiqui, 2020; Nelson & Campbell, 2017; Sharples, 2013).

There are two distinct forms of research evidence that educators are expected to engage with when developing and improving approaches to teaching. The first is evidence derived through teacher practitioner enquiry, an example of which is teachers’ engagement in lesson study and/or action research projects, and the use of pupil or school level data to aid decision making (see Brown & Greany, 2018). The second is accessing and applying the external research evidence, which consists of published peer-reviewed primary studies and evidence reviews (e.g., systematic reviews and meta-analyses), commonly produced by the research community. Over recent years, a number of research-engaged schools have been established in the United Kingdom (Godfrey, 2016; Nelson & Campbell, 2017; Sharples, 2013). A research engaged school according to Godfrey (2016, p. 7), promotes practitioner research amongst staff by encouraging staff to read and respond to published research; engages in research projects with other schools, agencies and higher education institutions; and, use research to inform decision-making. However, Godfrey and Brown (2018) argue that the ways in which this can be achieved have yet to be explicitly defined in the literature.

Currently there is very little evidence to suggest that using more evidence-informed methods in schools leads to greater impact on pupil outcomes (Cain, 2016; Gorard, See & Saddiqui, 2020). However, implementing programmes and methods that have an established causal-evidence-base are more likely to lead to positive effects (Goldacre, 2013; Slavin, 2020). Gorard et al. (2020) postulates that understanding what does not work also has its merits, as it prevents educators' time and school budgets being wasted, and allows limited resources to be distributed to provision that is more likely to work. Over recent years evidence repositories such as the Best Evidence Encyclopaedia (BEE), Evidence4impact, and the Education Endowment Foundation (EEF) have been created to provide teachers with better access to the research evidence. One of the core features of the EEF is the teaching and learning toolkit (https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/; Higgins et al., 2016), which provides educators with accessible summaries of effective teaching strategies and programmes.

Even though information about what works is now more available, it appears that there remains a lack of consistent use of evidence to inform teaching practice (Brown & Zhang, 2016; Coldwell et al., 2017). According to Ager and Pyle (2013), 67 per cent of school leaders indicate that they access evidence repositories when deciding how to spend their Pupil Premium funding (funding to support disadvantaged learners in England), and nearly half referred to using the EEF teaching and learning toolkit. However, the EEF (2018) found that 59 per cent of senior leaders were accessing the Teaching and Learning Toolkit to inform approaches to teaching. Vanderlinde and van Braak (2010) propose that teachers were more sceptical than school leaders about the value and relevance of educational research. Although, encouragingly, Brown and Greany (2018) reported that 86 per cent of teachers consider research to be useful when applying new approaches, and 76 per cent indicate that research informs their practice. Furthermore, recent research from Nelson et al. (2017) and Walker et al. (2019) shows that evidence repositories such as the EEF only had a small to moderate influence on decision-making in schools. Whilst the literature does suggest that educators are engaging with evidence, it does not demonstrate that decision-making, and importantly the provision that schools adopt into use, are underpinned by research evidence (Higgins, 2016). Some international empirical evidence reports that schools do not adopt evidence-based interventions (Gottfredson & Gottfredson, 2002; Hallfors & Godette, 2002). However, at the time of conducting the research for this project, there was very little evidence from the UK. The study conducted in Chapter 3 aimed to explore the range and evidence-base of interventions used across a typical cluster of schools in Wales, with the aim of better understanding the extent that research-evidence is in use in practice and, more specifically, how many interventions in use are evidence-based.

The barriers to an evidence-informed approach

To increase the use of evidence and the uptake of evidence-based provision in schools, it is important that policymakers and researchers identify the factors that enhance or hinder evidence use. There are many barriers to evidence use that have emerged in the last 20 years, and Hemsley-Brown and Sharp (2003) wrote one of the earlier reviews of the literature that illuminated these issues. Hemsley-Brown and Sharp identified that teachers had limited time to access research; lacked the research literacy skills required to understand it; and, believed the evidence was inaccessible, irrelevant, unhelpful, and often ambiguous. Since this review was published, other reviews and explorations a have been conducted and found that the same barriers persist (Cain, 2016; Gorard et al., 2020; Van Schaik et al., 2018; Walker et al., 2019).

One of the biggest barriers that teachers mention is the lack of time to source, understand, and apply evidence (Cain, 2016; Malin et al., 2020). Malin et al. (2020) and Van Schaik et al. (2018) report that as a result of diminishing budgets, schools often lack the structure required to facilitate an evidence-informed practice. Even if teachers had more time to dedicate to sourcing evidence, the literature is considered to be largely inaccessible to teachers (Van Schaik et al., 2018). Over recent years, more research evidence has become more readily available through evidence repositories such as the EEF. However, the majority of education research is published in online journals and remains inaccessible to teachers behind journal paywalls. Furthermore, the expanse of research literature across numerous national and international research websites, can make sourcing relevant research a challenge (Broekkamp & Van Hout-Wolters, 2007). According to Sparks (2018) many teachers are still unclear about the best places to access evidence.

If teachers are to effectively apply research findings in their everyday practice, then it needs to be usable and reliable (Van Schaik et al., 2018). More often research is written for the research community, and as such is viewed by teachers as complex and confusing. A large proportion of the teaching profession do not have the research literacy skills that are required to fully comprehend a research paper, which makes evaluating the evidence challenging, and time consuming (Hemsley-Brown & Sharp, 2008). Despite limited research literacy skills, teachers should be able to identify relevant evidence, draw reliable conclusions from the results, and apply the evidence to the classroom. However, some of the more popular programmes and teaching approaches have bodies of evidence of varying quality and reliability, which often result in contradictory findings. Consequently, teachers’ efforts to engage in and use research are often inconclusive, and sometimes lead to the uptake of programmes with weak causal-evidence (See, 2018).

Another important barrier to evidence-use is the lack of relevant evidence available (Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018). According to diffusion of innovations theory, the perceived compatibility of a programme is an essential attribute that facilitates decisions to adopt programmes. Consequently, educators require information about the characteristics of a programme, and how it functions (Rogers, 2003). However, according to Connelly et al. (2018) two thirds of RCTs conducted in education from 1980 to 2016 neglected to address questions of implementation and causal process and, as a result, teachers view the evidence-base as being of little utility (Van Schaik et al., 2018).

The facilitators of an evidence-informed approach

To support the use of evidence in schools it is suggested that school leaders can make positive changes to engage teachers (Coldwell et al., 2017). For example, to build a culture of evidence use, school leaders could schedule time for teachers to engage in research, provide research agendas, raise expectations, and demonstrate how research and evidence can inform school improvement (Brown & Zhang, 2016; Coldwell et al., 2017; Nelson & O’Beirne, 2014; Nutley et al., 2007).

In addition, researchers recommend that senior leaders and teachers should be provided with educator-friendly reviews of research that clearly and concisely present evidence for specific programme and methods (Gorard et al., 2020; Nelson & O’Beirne, 2014; Slavin, 2020). According to Williams & Coles (2007), teachers prefer research articles that are easy to read, and presented in more accessible formats (Williams & Coles, 2007). Nelson and O’Beirne (2014) state that it may not be enough to simply summarise the results. Bryman (2006) suggests that researchers need to ensure that research design and methods are transparent and limitations or threats to validity and reliability are clearly communicated within the summaries (Bryman, 2006). It is also suggest that intermediary organisations have an important role to play in translating and communicating evidence to schools (Nelson & Campbell, 2017; Nelson & O'Beirne, 2014). The second part to the study outlined in chapter 3 evaluates the impact of an educator-friendly review of evidence on the provision that schools adopt.

To contribute to the development of teachers that are already working in the profession, in-service education and training (INSET) days are held in schools to cover a wide range of training activities (see Craft, 1996; Hargreaves & Fullan, 1992). To increase teachers' skills and knowledge, research indicates that in-service training should be conducted on concepts and ideas surrounding evidence-into-action (Hemsley-Brown & Sharp, 2003; Nelson & O'Beirne, 2014). Klingner, Boardman, and McMaster (2013) indicate that teachers are more likely to adopt evidence-based programmes and put them into use if they have attended a training course. Although training is described as an effective way of developing teachers' knowledge and skills, evidence also suggests that it may not be enough to result in meaningful changes to practice (EEF, 2019). For meaningful change to take place, adequate re-training and coaching of staff is required to promote mastery, motivation and sense of self-efficacy (Dusenbury & Weissberg, 2017; Durlak & DuPre, 2008; Ransford et al., 2009). The study outlined in Chapter 5 explores the effects of a typical in-service training day on the uptake and implementation of the Good Behavior Game.

To create a relevant and useful evidence base for schools, researchers need to provide answers to the questions that are most relevant to teachers and school leaders. Arguably, researchers should engage in greater use of mixed methods research to achieve this, plus a consideration of a question-to-methods approach when planning research (Owen et al., 2022) For example, a teacher should be able to discern from the evidence if and how a programme will work in their context and with the available resources (Nutley et al., 2013). Schaik et al.

(2018) and Gorard et al. (2020) report that greater collaborations and partnerships between researchers and teachers are necessary to facilitate the creation of evidence that is more applicable for practice (Gorard et al., 2020; Hemsley-Brown & Sharp, 2008; Schaik et al., 2018).

Adoption

To advance the uptake of evidence-based methods in schools, it is important that policymakers and researchers identify the factors that facilitate decision-making and the adoption process. Recent survey-based research conducted in the UK by Nelson et al. (2017) and Walker et al. (2019) suggests that decisions about approaches to teaching are more often informed by past experience of what works, and the experiences of teachers within school and in other schools. Whilst the findings from Nelson et al. (2017) and Walker et al. (2019) are useful, they primarily explore whether schools engage in meaningful evidence use. To address the lack of detail, Chapter 4 utilised interviews to explore in greater detail the factors that inform school leaders' choice of teaching programmes in schools. However, other findings from the USA indicate that educators may reject evidence-based methods if they perceive them to conflict with a school's policies, philosophies, values, existing provision, needs, and school resources (Bumbarger, & Perkins, 2008; Forman, et al., 2009; Powers, Bowen, & Bowen, 2010). Similarly, Greany and Brown (2017) report that programmes and methods considered popular and compatible with teaching practice are more likely to be adopted into use despite them having little or no evidence demonstrating effectiveness.

Diffusion of innovations theory

The factors that facilitate adoption-decisions in schools outlined above, closely align with a widely used adoption model proposed by Rogers (2003). Diffusion of innovations theory, defines how new ideas or practices (innovations) spread (diffuse) between members of a social system. According to Rogers, decisions to adopt new ideas and practices within an organisation are largely influenced by communications between members of social networks.

Rogers refers to members of social networks as ‘change agents’. Within the context of an education system, change agents are commonly staff in schools (internal networks), but can also be members of external networks that are working towards a common goal. For example, an external change agent might be a local school improvement advisor and/or a researcher from a local university. The decisions of whether to adopt, implement and sustain use of an innovation move through five phases. The first three phases are considered to be the pre-implementation or adoption stage, and within this stage decisions of whether an innovation will be adopted, take place. *Knowledge* – an adopter becomes aware of an innovation and how it functions; *persuasion* – an attitude is formed towards the innovation; and, *decision* – adopter chooses whether to adopt or reject an innovation. The latter two phases are considered to be the implementation stage, and within this stage adopters make decisions about the extent that an innovation will be implemented and maintained. *Implementation* – an innovation is put into use, and may be modified by the user; and, *confirmation* – use of the innovation is sustained if decisions to adopt are reinforced (Rogers, 2003).

According to Rogers, and a recent review of the diffusion literature by Wisdom et al. (2014), the process of adoption is affected by a range of other factors, such as the perceived characteristics of an innovation and how it functions. Five characteristics have been proposed that facilitate adoption, and these are: *relative advantage* (degree to which an innovation is perceived as better than the existing one), *complexity* (degree to which the innovation is perceived as difficult to use and understand), *compatibility* (degree to which an innovation is compatible with existing needs and values), *trial-ability* (degree to which the innovation can be tested), and *observability* (degree to which the innovation can be viewed in action).

Implementation

Although not always assessed, the evidence suggests that when programmes are adopted into use in schools, implementation is often inconsistent, and this has been found to compromise effects (Durlak & DuPre, 2008; Domitrovich & Greenberg, 2000; Durlak et al., 2011). The five components of implementation that can be evaluated through self-report measures and observations are: *quality* (the manner and level of skill in which a programme is delivered), *fidelity* (the degree to which the essential key components of a programme are maintained), *participant responsiveness* (the extent that participants are engaged and respond to a programme) *dosage* (frequency of delivery), and *adaption* (the extent to which a programme is change or modified) (Fixen et al., 2009).

Evidence suggests that factors such as a limited resources (time & money), incompatibility of programmes, lack of drive from leadership, lack of training and coaching, and impact from competing priorities and other programmes have been found to prompt adaptions to programmes (Domitrovich et al., 2008; Durlak, 2016; Rogers, 2003). Cain (2016) finds that teachers often view their context as highly complex and disparate to other school contexts and, therefore, where such conditions exist it is unlikely that a programme will be implemented fully and adaptions to a programme are inevitable. Evidence from the implementation literature reports that universal SEL programmes implemented with low levels of fidelity are less likely to have a positive impact on pupils' social and emotional skills, and similar outcomes are reported for low levels of implementation quality (Durlak et al., 2011; Greenberg et al., 2001). However, the majority of implementation findings report positive effects from programmes even when optimal dosage is not achieved (Domitovich & Greenberg, 2000; Durlak & DuPre, 2008).

Durlak and DuPre (2008) report that implementation fidelity rarely achieves 80 per cent, but despite lower levels of implementation fidelity, positive effects can still be achieved, suggesting that there is room for adaption across most programmes. Yet, more often, evaluations of programme implementation neglect to focus on adaptions and only measure fidelity and dosage, and sometimes quality (Durlak & DePre, 2008). An understanding of the elements of programmes that can be adapted might mean that a greater number of evidence-based programmes are perceived as being compatible by a greater number of educators. In chapter 6 we describe the process of contextual adapting PATHS, and evaluate the impact and implementation of the newly adapted programme.

Welsh education system and Curriculum for Wales

Wales is a bilingual country with 440 Welsh medium schools, and the teaching of the Welsh language is compulsory for all pupils up to the age of 16. There are approximately 470,244 pupils in 1,470 local authority maintained schools in Wales. Across these schools there are 74,595 (15.9%) pupils with additional learning or special educational needs, and 380,139 (23.6%) pupils aged 5 to 15 that are known to be eligible for free school meals (Welsh Government, 2021). Even before the coronavirus pandemic, 3 in 10 children in Wales were known to be living in poverty (Joseph Rowntree Foundation, 2020). The detrimental effects of child poverty on educational attainment and on pupil mental health and emotional wellbeing is well documented (Bradshaw & Main, 2016; Egan, 2007; Van Lancker & Parolin, 2020; Yoshikawa & Beardslee, 2012). In an attempt to reduce the impact of child poverty and improve outcomes for disadvantaged learners, the Welsh Government introduced the Pupil Development Grant (PDG), which provides additional funding for pupils eligible for free school meals (eFSM) and for Looked After Children (LAC), to provide additional instruction and/or support through school-based interventions (Welsh Government, 2015).

Since the creation of the Welsh assembly in 1999, The Welsh Government's Department for Education (DfE) has been responsible for education policy in Wales. Following disappointing performance in the Programme for International Student Assessment (PISA) in 2009, the Welsh Government commissioned the OECD (2014) to conduct an independent review of the Welsh education system, which was the catalyst for a number of significant policy changes. To build professional capital and learning within the Welsh education system, the OECD (2014) recommended the need for a more collaborative and research-informed approach. *Qualified for Life: An Education Improvement Plan* (Welsh Government, 2014), and *Education in Wales: Our National Mission* (Welsh Government, 2017), outlined large-scale initiatives to enhance the quality and equity of education through a self-improving system. Within this system education professionals collaborate within school and between schools, and with other organisations with a vested interest in improving educational outcomes, to develop and share best practice, based on the evidence of what works.

In 2013, the Welsh Government created four regional consortia and transferred the responsibility for school improvement services from the local authorities to the newly established consortia (Welsh Government, 2015). The Regional School Effectiveness & Improvement Service for North Wales (GwE) is the consortia responsible for providing school improvement services to 401 maintained schools across six local authorities in North Wales. Through the work of school improvement advisers, GwE supports schools to improve school provision, leadership and learner outcomes. To develop a more collaborative way of working in line with the recommendations from the OECD (2014) and the strategic objectives of Welsh Government (2014, 2017), GwE and the Collaborative Institute for Education Research, Evidence and Impact (CIEREI) at Bangor University formed a strategic research partnership that aimed to build research capacity within North Wales schools. Born within this collaboration was a ‘collaborative model of working’ that involved university researchers and GwE working in partnership with schools across the region co-constructing close-to-practice research that aligns with the priorities of the organisations involved, and addresses the issues pertinent to the schools (Tyler et al., 2019).

The empirical research presented in Chapters 3, 4, 5 and 6 were conducted with the staff and pupils from a well-established cluster of schools in Wales, which comprise two secondary schools (English medium), seven primary schools (six English medium and one Welsh medium) and one special school (English medium). The cluster entered into a collaborative research partnership with GwE and CIEREI, Bangor University, to commission close-to-practice research that aligns with the values and priorities of the organisations involved, and importantly addresses the issues pertinent to the schools. The aim of the collaborative partnership was to co-construct research that would move the cluster towards a more evidence-informed practice. A priority for the cluster senior leaders was to review the evidence base for the interventions already in use across schools, and the area of focus most pertinent to the schools was improving the wellbeing of pupils. Throughout the project, the head teachers and senior leaders were closely involved in the development and planning of research and worked collaboratively with the lead researcher who was embedded within the cluster. From a research perspective, we aimed to explore ways to increase the adoption and use of evidence-based methods in schools.

Following an independent review of the curriculum in Wales led by Graham Donaldson, the subsequent Successful Futures report (Donaldson, 2015) set out to radically restructure the school curriculum (Welsh Government, 2017). Within the new Curriculum for Wales (CfW) schools are responsible for the realisation of a more relevant and engaging local curriculum driven by four main purposes, which are the shared vision and aspiration for all pupil across Wales to be ambitious, capable learners, ready to learn throughout their lives; enterprising, creative contributors, ready to play a full part in life and work; ethical, informed citizens of Wales and the world; and, healthy, confident individuals, ready to lead fulfilling lives as valued members of society. The new CfW is centred on six Areas of Learning and Experience (AoLE) that aim to develop learners both academically, and socially and emotionally, and these are: expressive arts, health and well-being, humanities, languages, literacy and communication, mathematics and numeracy, and science and technology.

There is now wide agreement from educators, researchers and policy makers that schools have a moral imperative to enhance the social and emotional skills of pupils in schools, which is reflected in the new CfW (Donaldson, 2015; Payton et al., 2000; Payton et al., 2008; Welsh Government, 2017). The health and wellbeing (HWB) AoLE, which replaces personal and social education, encompasses a whole-school approach that aims to promote physical health and development, mental health, and emotional and social well-being of all learners in schools across Wales.

Social and emotional approaches in schools

Social and emotional competence is the ability to recognise and manage emotions, resolve problems effectively, show empathy, set and achieve goals, and establish and maintain positive relationships (Collaborative for Social and Emotional Learning [CASEL], 2015). There is an expansive literature base that advocates for the teaching of whole-school (universal) social and emotional skills to develop pupils' social-emotional competence to promote healthy behaviours and to prevent behavioural problems occurring (e.g., CASEL 2002; Elias et al., 1997; Greenberg et al. 2003; Zins & Elias, 2007). Through social and emotional learning the following core competencies are developed: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (CASEL, 2015). There are now many evidence-based social and emotional programmes available to schools that support teachers to develop pupils' social and emotional skills (see Corcoran et al., 2018; Durlak et al., 2011; Greenberg, Domitrovich, & Bumbarger, 2001; Murano, Sawyer, & Lipnevich, 2020).

Two examples of evidence-based teaching strategies that support the development of pupils' social and emotional competencies are the Good Behaviour Game (GBG) and the Promoting Alternative Thinking Strategies (PATHS) programme, and both of these are the focus of the research conducted in Chapters 5 and 6.

Good Behavior Game (GBG)

The Good Behavior Game (GBG) is a classroom-wide procedure that was developed by Barrish, Saunders, and Wolf (1969) to reduce the disruptive behaviours (out-of-seat and talking-out behaviours) of students in elementary classrooms (ages 5-10 years). The procedure has a strong theoretical base and extensive evidence supporting its use in improving behaviour across a range of school contexts and with a variety of populations (see Bowman-Perrott et al., 2016; Coombes et al., 2016; Flower et al., 2014; Groves & Austin, 2017; Joselyn et al., 2019; Kellam et al., 2011; Nolan et al., 2014; Wiskow et al., 2019), and has been described by Embry (2002) as cost-effective and simplistic. Much of the evidence supporting its use has been focused on decreasing disruptive behaviour with primary (elementary) aged pupils, although it has also been used successfully with secondary school age students. Since its development, multiple group design and single-case design studies conducted nationally and internationally have shown the GBG to be effective in reducing challenging and disruptive behaviours and increasing pro-social behaviours in the classroom (e.g., Bowman-Perrott et al., 2015; Coombes et al., 2016; Flowers et al., 2014). However, recent evidence from a large-scale EEF trial, conducted in the UK, reported more mixed findings (Humphrey et al., 2018). Humphrey et al. found no evidence that the GBG was effective, but they did find tentative evidence that boys at-risk of developing conduct problems benefitted from the game. To date, the strategy has been included in a number of research repositories and registries such as: the Early Intervention Foundation's (EIF) Guidebook (http://guidebook.eif.org.uk/), the Substance Abuse and Mental Health Services Administration (SAMHSA): National Registry of Evidence-based Programs and Practices (https://www.samhsa.gov/resource-search/ebp); and has been certified as a 'promising' programme by the Blueprints for Healthy Youth Development database (http://blueprintsprograms.com/).

The interdependent group contingency procedure employs a number of behavioural management techniques such as the teaching of classroom rules or expectations; the acknowledgment of prosocial behaviours; and the reinforcement of behaviours through praise, reward, and feedback (Flowers et al., 2014). To play the game the following procedure is followed (1) divide the class into teams; (2) remind pupils of the rules; (3) announce the start of the game and set a timer; (4) award points and provide praise to the teams who follow the rules; and, (5) announce the winning team(s) and provide a reward (s). The study outlined in chapter 5 explores the effects of a typical in-service training day on the adoption and implementation of the GBG.

Promoting Alternative Thinking Strategies (PATHS)

PATHS is a school-based universal social and emotional learning (SEL) programme that was developed in the USA by Greenberg and Kusche (1993). The PATHS® programme is based on the Affective-Behavioral-Cognitive-Developmental model of development (ABCD). To promote social and emotional competence the model focuses on the developmental integration of affect, emotion language, behaviour and cognitive understanding (Greenberg & Kusche, 1993). The core components of the programme include a taught curriculum and generalisation activities. Teaching resources include lesson plans and activities that cover topics such as identifying and labelling feelings, controlling emotions, and understanding others perspectives.

PATHS has a strong theoretical foundation and is supported by an extensive international literature base, which includes multiple randomised controlled trials (RCTs) (e.g., Domitrovich, Cortes, & Greenberg, 2007; Greenberg et al., 1995). PATHS is more commonly used as a whole school or universal approach to teaching social and emotional skills and the evidence has predominately been conducted in the general education setting. However, research has also been conducted with special education populations and found that the positive effects can be replicated (Kam et al., 2004; Greenberg et al., 1995). In the USA, PATHS is defined as a ‘model program’ by the Center for Study and Prevention of Violence (CSPV, 2006) and the Substance Abuse and Mental Health Services Administration (SAMHSA, 2011).

The PATHS programme has been adapted in many diverse English and non-English speaking cultures nationally and internationally. In fact, the PATHS materials were culturally adapted by children’s charity Barnardos, for use in the UK. Following an influential government report PATHS UK was recommended for widespread adoption based on its promising evidence-base (Allen, 2011). Nevertheless, the findings from research conducted in the UK are more mixed. For example, evidence from two quasi-experimental study reported significant and positive effects on a range of social and emotional competencies and mental health outcomes for pupils that used the pre-school PATHS (Curtis & Norgate, 2007; Hughes & Cline, 2014), and a large-scale longitudinal RCT conducted in Northern Ireland found small positive effects of PATHS on social and emotional competencies (Ross et al., 2011). However, a more recent large-scale cluster-RCT conducted in England found no statistically significant differences between participants in the PATHS group compared to the control on measures of mental health (Berry et al., 2015), and most recently, evidence from a large-scale cluster-RCT in the UK indicated that PATHS had small but significant improvements on pupils’ social skills and psychological wellbeing, and no impact on academic outcomes (Humphrey et al., 2018). However, an economic analysis of PATHS suggest that it has the potential to be cost-effective when used to promote social and emotional skills (Turner et al., 2020).

The results from the PATHS implementation literature report variable levels of implementation, which are on the whole are associated with pupils outcomes. Nevertheless, often reports of implementation fidelity are relatively high, but PATHS is rarely implemented at the frequency recommended by developers. Most of the PATHS literature report lower levels of dosage, commonly around 50 per cent. However, less than optimal dosage is not associated with greater negative outcomes (Barlow & Lendrum, 2018; Berry et al., 2015; Faria et al., 2013; Hennesey & Humphrey, 2020; Humphrey et al., 2016), and suggests it is possible for teaching staff to adapt the programme and achieve good outcomes. Qualitative interviews with teachers report that some view PATHS UK as aligning with values and needs. However, issues have been raised regarding the appropriateness of the language and cultural references used in the materials, and the length of lessons. Teachers commonly report difficulties finding adequate time and space in the curriculum to deliver PATHS at the frequency required. As expected, teacher adaptions to the content and materials of PATHS have been reported (Faria et al., 2013; Humphrey et al., 2016). Chapter 6 outlines the process of adapting PATHS and includes a feasibility evaluation.

Overview of thesis

This PhD was partially funded by the school cluster and GwE. The primary focus of the following empirical chapters were to explore methods to increase the adoption of evidence-based methods in the school cluster, and to inform the wider knowledge-base on how to increase the use of evidence-based programmes in schools.

Chapter 2 provides context and justification with regards to the research methods used throughout this thesis.

Little is known about the range of intervention programmes and approaches used in UK schools, or the evidence supporting them. Some international evidence suggests schools do not adopt evidence-based interventions (Gottfredson & Gottfredson, 2002). **Chapter 3** is a mixed-methods study that consists of a questionnaire and a rapid systematic review that were used to provide evidence on the range and evidence-base of interventions used in the school cluster. The findings indicate that schools adopt a large number of interventions that have very little causal evidence supporting them, and few have robust evidence of causal-effect. The findings from a follow-up study using interviews, suggests that when presented with summary reports of the evidence, not all schools use the information to make evidence-informed decisions about provision. Chapter 3 has been published by BERA in Review of Education journal, and in the BERA blog (see https://www.bera.ac.uk/blog).

In light of the findings from the evaluation of cluster provision in Chapter 3, we were interested to explore the factors that influence the head teachers' decision-making when choosing their school provision. To date, there has been a number of survey-based studies conducted on this topic, which primarily aimed to explore teachers' use of research evidence (Nelson et al., 2017; Walker et al., 2019). The findings from these studies reported that decision-making is informed by the opinions and experiences of staff in schools.

Using qualitative interviews, **Chapter 4** explores in greater depth the factors that influence head teachers' decisions when choosing which teaching programmes to adopt. We chose to focus the interview questions to specific programmes that the head teachers' had recently adopted, to gain a realistic insight into the variety of factors that inform adoption-decisions. The most pertinent themes derived from our data support theory and previous survey findings, and expand them (Nelson et al., 2017; Rogers, 2003; Walker et al., 2019). Similarly, we found that the advice and experience from staff in other schools, particularly other senior leaders, was frequently sought and highly trusted and valued. In addition, we discovered that it was common practice for senior leaders to visit other schools to observe a programme of interest in use to determine the compatibility of the programme and how it functions. Importantly, we found that information gathered from other schools often influenced decisions to adopt or reject programmes. Additionally, we found that senior leaders would seek recommendations about evidence-based methods from their local school improvement advisor (GwE), who was viewed as a trusted source of knowledge for research and evidence-based practices. Finally, we found that the compatibility of programmes was a critical factor in the adoption-decision process, and that often head teachers would seek programmes that placed minimal demand on resources (staff time and money); were time saving; they could be delivered by computers or digital platforms; and could be accessed at home. Chapter 4 contains a draft manuscript that we have submitted to The Wales Journal of Education and is currently under review.

Evidence suggests that there is a lack of consistent use of evidence to inform teaching practice (Brown & Zhang, 2016; Coldwell et al., 2017). Based on findings from Chapter 3 and Chapter 4 we also found that research evidence is infrequently used to inform the programmes that schools use. The GBG is a promising strategy that has a wealth of evidence supporting its effective use in the classroom (Flowers et al., 2014). Nevertheless, it appears it has failed to translate into every-day use in the classroom (Gottfredson & Gottfredson, 2002). It is recommended that in-service training should be used to increase teachers' skills and knowledge of evidence-based methods (Hemsley-Brown & Sharp, 2003; Nelson & O'Beirne, 2014). Klingner et al. (2013) suggests that teachers are more likely to adopt evidence-based programmes and put them into use if trained. With the aim of increasing the use of evidence-based provision in the cluster of schools, we trained teachers in the use of the GBG during a typical in-service training day. With the exception of staff time, there was no impact to the schools resources, and they were under no obligation to adopt and use the GBG.

The aim of the study in **Chapter 5** was to explore the extent that teachers implement the GBG following typical in-service training, and the factors that facilitate decision to adopt or reject. The findings suggest that after completing in-service training (February) all of the teachers had good knowledge of the GBG. We found that six teachers that played the GBG in the classroom, viewed it as an acceptable procedure. However, only two teachers chose to adopt the GBG, three teachers rejected it immediately after training, four teachers trialled it and then chose to reject, and for two teachers their decisions were unknown. Furthermore, one teacher discontinued use of the GBG at time-point 2 (March and April), and by time-point 3 (July) none of the teachers continued to implement the GBG. Observation data from one adopting teacher at time-point two, revealed that implementation fidelity was high. However, interview data revealed that the same teacher played the game infrequently. The most pertinent data from the interviews revealed that four teachers rejected the GBG because they perceived other competing classroom management strategies to be more advantageous, or too similar to warrant change (relative advantage); and three teachers rejected the procedure because they viewed it as incompatibility (compatibility) with class needs, in particular the time of year in which training was conducted, was not suitable. Chapter 5 is a draft manuscript that we intend to submit for publication.

Evidence from Chapters 4, 5 and across the implementation literature suggests that the perceived compatibility of a programme impacts adoption-decisions (Rogers, 2003) and the extent that a programme is implemented (Durlak & Dupre, 2008). Through anecdotal discussions with staff from the school cluster, we discovered that PATHS has been adopted into use by the primary schools, and subsequently rejected. In alignment with findings from Humphreys et al. (2018) the teachers reported that the programme was insufficiently anglicised, complex, and the time required to deliver lessons was too long. As a result, and based on the promising evidence-base for PATHS (e.g., Domitrovich et al., 2007; Greenberg et al., 1995), we agreed with the senior leaders that for the final study, we would work with staff to adapt PATHS, to create a more cultural and contextual relevant programme. Through feasibility research we aimed to evaluate, if the adapted PATHS programme would lead to adoption, high levels of implementation, positive attitudes, intentions to sustain use, and positive impact on pupils' social and emotional competencies.

Chapter 6 describes the process of adapting PATHS, and includes a feasibility evaluation of the adapted programme that utilises: structured observations and teacher-report implementation checklists to evaluate multiple components of implementation; focus groups to address suitability, acceptability, and further adaptions; and, a pre- and post-test quasi-experimental design to measure the impact of the adapted programme on Year 1 pupils' social and emotional competencies. Our results revealed that the adaption objectives were achieved, teachers attitudes towards the adapted PATHS programme were overall positive, levels of implementation were high, and evidence of medium-large causal-effects on pupil social-emotional competencies were observed. Thematic analysis of the data revealed that competing priorities had a negative impact on dosage, and competing programmes may prevent sustained use. Chapter 6 is a draft manuscript that we intend to submit for publication.

Chapter 7 is the general discussion, which outlines and discusses the findings from the empirical chapters, the strengths and limitations, and implications that arise.

Chapter 2: Methods and design

We adopted a pragmatic approach (Coe, 2012) to the empirical research outlined in Chapters 3, 4, 5 & 6, and chose the most appropriate research methods to best answer the research questions that we aimed to address. In line with the evidence building framework by Owen et al. (2022), this would be considered a “question-to-methods” approach. For Chapter 3, 5 and 6 we adopted a complementary mixed-methods design, which according to Gray (2018) is an approach that consists of combining both qualitative and quantitative methods within a study to measure overlapping constructs and to address the different research questions posed. Pring (2000) argues that both quantitative and qualitative research methods are important in educational research to capture the complexity of school contexts. Through adopting a mixed-methods approach, a clearer and deeper understanding of a phenomenon can be established (Ponce & Pagán-Maldonado, 2015).

In order to answer the research questions in the study outlined in Chapter 3 we used a mixed methods design incorporating a questionnaire, systematic review and qualitative interviews. As very little is known about the interventions used in UK schools, we developed a bespoke questionnaire that included a broad range of question items designed to capture the wide range of provision in schools. To understand the range of interventions used across the school cluster, we used an open-response questionnaire that asked school leaders to provide a list of all school-based interventions used in the current academic year. A thorough search of the literature revealed that a widely accepted single definition for the term ‘intervention’ in educational research did not exist. We therefore chose to develop an inclusion and exclusion criteria for an intervention derived from the meaning of tier two and tier three intervention used in the Response-to-Intervention (RtI) approach (Fletcher & Vaughn, 2009).

RtI is an empirically derived system of support, developed in the USA to provide research-based instruction and interventions to students in schools. The approach responds to deficits in learning and/or wellbeing/behavioural issues by using evidence-based intervention with different tiers of intensity to meet pupils’ needs. Tier one is considered to be typical whole school / universal provision; tier two is short to medium term intervention, which is delivered through small groups to approximately 15 per cent of pupils; and the third tier of intervention involves intensive individualised support for pupils with persistent and significant need, which is includes approximately five per cent of the school population. (Fletcher & Vaughn, 2009). Based on the RtI definition of an intervention, we collected data using an iterative approach at three time-points to ensure the final data set was comprehensive, and that any newly implemented interventions were not missed. As described by Kerssens-van Drongelen (2001), an “iterative theory-building process” allows for the modification of the research questions over time based on the data collected.

To assess the causal-evidence for the interventions collected in Chapter 3, we chose to adopt a rapid review approach. “A rapid review is a type of knowledge synthesis in which components of the systematic review process are simplified or omitted to produce information in a shorter period of time” (Khangura et al., 2012, p.10). Khangura et al. suggest that a rapid review is suitable when reviewing a new or emerging topic, and/or the research question is broad. As little was known about the interventions in use, and we were searching across a broad range of topics (e.g., literacy, numeracy, and social-emotional learning programmes), a rapid review was deemed the most suitable approach. Another necessary component of a typical systematic review, which is not a requirement of a rapid review, is to have an independent person review the articles for bias (Khangura et al., 2012). As a result of budget and time constraints of a PhD project we were not able to employ an independent reviewer to carry out an independent screening of the studies.

When synthesising or reviewing a body of evidence, it is important to critically examine the design and execution of research to determine the trustworthiness of findings and understanding what counts as high-quality and low-quality evidence (Gorard et al., 2020). It is common for reviews of research to neglect to attend to the quality and commensurability of studies, when reviewing or aggregating the results, which can lead to untrustworthy research findings and misleading and often inconclusive findings (Gorard et al., 2020). There are many systems used for evaluating the quality of research, and for the review of evidence conducted in Chapter 3, we used the sieve method to judge the quality and trustworthiness of the included studies (Gorard, 2014). The benefit of the sieve method is it can be used to judge the quality of findings for any type of research (Gorard et al., 2020). Each study included in this review was given an individual rating from 0🔒, which represents a study that adds nothing to the knowledge base; to 4🔒, which represents a trustworthy and robustly designed study. The rating is based on five criteria: design, scale, missing data, data quality, and other threats to validity.

A qualitative approach was adopted to explore the senior leaders' decision-making process when adopting academic and wellbeing programmes (Chapter 4). Survey research has previously been conducted with teachers to evaluate the extent that teachers use research evidence to inform practice (Nelson et al., 2017; Walker et al., 2017). Within these surveys Likert scale response questions have explored the other sources of information that influence decisions about teaching and learning. However, we felt that interviews were also necessary to explore in greater detail the experiences of senior leaders when making decisions about school provision, and to gain a deeper understanding of the topic (Rubin & Rubin, 1995). We chose to adopt the Reflexive Thematic Approach (RTA) outlined by Braun and Clarke (2013, 2019, 2021). This approach was chosen because it acknowledges the impact of the researcher on the analytic process and on the creation of knowledge. Unlike other structured approaches to thematic analysis, the RTA process requires a more organic and recursive coding approach, and advocates for deep reflection on, and engagement with, data (Braun & Clarke, 2019). We adopted a predominantly inductive approach to best reflect the data. However, a degree of deductive analysis ensured that the coding contributed to producing themes relevant to the research questions, and allowed us to draw on the data from previous survey research (e.g., Nelson et al, 2017; Walker et al., 2019). In doing so, we were able to triangulate our data with existing survey findings to strengthen their validity (Denzin, 1978; see Chapter 4 for more detail on the specific approach used).

To understand the extent that the GBG was implemented following in-service training in Chapter 5, we collected quantitative data through structured observations that aimed to measure implementation fidelity. Arguably, a structured observation approach requires fewer assumptions and inferences, is less susceptible to confounding influences, and is less likely to result in biased results that are inherent with self-report measures of implementation (Heath et al., 2010). To collect more reliable and valid data, we used an existing GBG implementation fidelity rubric (Schaffer et al. 2006). Wiggins (1998) suggests that there are several benefits of using validated rubrics. One widely cited effect is the increased consistency of observations across the participants and the likelihood of collecting more accurate data, which is representative of the variable being measured.

To better understand adoption and the factors that facilitated decisions to adopt the GBG, we chose to conduct semi-structured interviews. The themes derived from the interviews conducted in Chapter 4 closely aligned with the concepts of Rogers (2003) diffusion of innovations theory. As a result, we decided to create a semi-structured interview guide that included open-ended questions and prompts, which aligned with the key constructs of Rogers' (2003) theoretical framework. We analysed the interview data using a deductive approach to content analysis, which Hsieh and Shannon (2005) and Newby (2010) refer to as 'directed' content analysis. This approach was chosen because of its ability to systematically search for the presence of, and code for, certain words, or concepts within a dataset, which are derived from a theoretical framework (Hsieh & Shannon, 2005).

In Chapter 6, we culturally and contextually adapted the PATHS programme for the cluster schools and evaluated the newly adapted programme using feasibility research. Overall, the original PATHS programme has an extensive evidence base that demonstrates positive impact on pupils’ social and emotional outcomes, but the newly adapted PATHS programme lacks evidence. According to Bowen et al. (2009), feasibility studies are a cheaper and simpler approach, which can be used to produce findings that help determine whether a programme should be recommended for further testing through a more robustly designed efficacy study.

To evaluate the extent that PATHS was implemented we chose to measure implementation fidelity, quality, pupil responsiveness, dosage and adaptions through both teacher-report implementation rubrics and observations. Mitchel (1986) suggests using multi-methodologic triangulation to decrease the biases that stem from using a single method. In Chapter 6 we chose to conduct observations to mitigate any effects from response-bias that is known to occur when using self-report measures (Van de Mortel, 2008).

To evaluate the impact of the adapted PATHS programme on pupil outcomes we used a pre- post-test quasi-experimental design. As is often the case with school-based research, it was not possible to assign pupils randomly to intervention and control conditions. Instead, we opted to use a non-probability sample and assigned two Year 1 classrooms in each school to either the intervention or the control group. However, there is no issue of statistical generalisation as the outcomes are only relevant to the population studied, and this was intended as pilot research only. In alignment with the rationale and statistical methods used by Gorard, Siddiqui and See (2014), we chose to report change scores and effect sizes to represent the difference in scores from pre- to post-test. The post-test scores and difference scores were averaged for each group (intervention and control), and the difference between the average scores were expressed as an ‘effect’ size. To determine the magnitude of effect we calculated Hedges' g by dividing the difference by the pooled standard deviation of the gain scores for both groups. We chose to calculate Hedge's g, because it uses pooled weighted standard deviations, which accounts for the imbalance between intervention and control groups (Hedges, 2008). To interpret the size of effects in Chapter 6, we used the thresholds outlined by Hedge's (2008), which closely align with the default approach to evaluating the magnitude of effect sizes proposed by Cohen (0.2 = small, 0.5 = medium, 0.8 = large; Cohen, 1969). Recent meta-analyses of robustly designed studies have found that education interventions often report small or null effect when applying Cohen's benchmarks (Cheung & Slavin, 2016). Kraft postulates that the persistent use of Cohens d to interpret effect sizes often results in the misinterpretation of the magnitude of effect. Kraft (2020) proposes a set of revised benchmarks for interpreting research outcomes for educational interventions for preschool to Key Stage 5, which attempts to draw on the practical considerations and contextual factors of a study. Accordingly, we also chose to interpret the effect sizes in Chapter 6 using these benchmarks.

At pre- and post-test teachers completed three teacher-report measures of pupils' social and emotional competencies. These measures were chosen because all three had been used in previous PATHS research and had high levels of internal consistency and reliability.

Lastly, we prepared a semi-structured interview guide that included open-ended question that aimed to explore the suitability and acceptability of the adapted PATHS programme. We analysed data using a thematic approach as outlined by Braun and Clarke (2006), and adopted a hybrid approach to the process of identifying and labelling themes that utilised both deductive and inductive analysis of the data (Fereday & Muir-Cochrane, 2006). This enabled our analysis to be informed by existing theories and evidence from the PATHS implementation literature, and ensured that emergent themes were included (Fereday & Muir-Cochrane, 2006).

Chapter 3: Assessing the Range and Evidence-Base of Interventions in a Cluster of Schools.

Pegram, J., Watkins, R. C., Hoerger, M., & Hughes, J. C. (2022). Assessing the range and evidence-base of interventions in a cluster of schools. Review of Education, 10(1), e3336.

Abstract

The last decade has seen an increased focus through policy and research for schools to move towards an evidence-informed practice. Although some practitioners now access the external research evidence when deciding which interventions to adopt in their school, research suggests many still do not. Instead, approaches to teaching and learning are often informed by trends and the opinions and experiences of practitioners. Little is known about what intervention programmes/approaches are used in schools and whether they are evidence-based. We conducted this study to assess the range and evidence of interventions used in a school cluster in Wales, which comprised two secondary schools, seven primary schools and one special school. Using questionnaires, we evidenced 242 interventions. Following screening, we included 138 of these in the analysis and categorised them according to the 'SEN Areas of Need'. We then conducted a rapid systematic review of the literature for these interventions and found that 30 per cent had some evidence of positive causal-impact on pupil outcomes, 67 per cent had no published evidence, and three per cent had causal evidence to suggest they were ineffective. One year later, we conducted a follow-up study to assess if schools' knowledge of the evidence for the interventions, presented through a summary report, had an impact on their provision. Our data suggest that the reports had very little impact on existing provision, and some schools continued to use the same interventions. The limitations of this study and directions for the cluster, policy and research are discussed.

The last decade has seen significant changes in national education policy designed to promote more evidence-informed approaches to teaching and learning in schools (Coldwell et al., 2017; Donaldson, 2015; Department for Education [DfE], 2010, 2016a, 2016b; Institute for Effective Education, 2019; Organisation for Economic Co-operation and Development [OECD], 2014; 2017). An example of this is the development of the self-improving education system in England (Hargreaves, 2010), and the promotion of schools as learning organisations in Wales (Kools et al., 2018). These policies were introduced as a means to developing best practice and improving standards of teaching, through the use of external research and teacher enquiry. Within these models teachers are viewed as ‘knowledge workers’, actively building their knowledge, expertise and capacity, through enquiry and collaborating within school, with other schools, and the wider research community (see Brown & Greany, 2018; DfE, 2010, 2016a; Kools et al., 2018, 2020; Welsh Government, 2015a, 2015b). The intended outcome is an education system in which schools and teachers are closely aligned with research findings and evidence, and the use of research in schools to improve provision is a cultural norm (Brown & Greany, 2018; Godfrey, 2014, 2017; Godfrey & Handscomb, 2019). However, the ways in which this can be achieved have not yet been clearly defined in the literature (Godfrey & Brown 2018).

There is a growing impetus for schools to access and apply external research evidence to improve the quality of teaching and the outcomes learners achieve (Gorard, 2020; Slavin, 2020). In this study, we consider external research evidence to be evidence derived from the external research community, and more often created by academic researchers. This evidence includes peer-reviewed and published primary studies, and reviews such as meta-analyses or evidence syntheses. The rationale for using research evidence to improve the quality of provision in schools has attracted steady support over recent years (e.g. Goldacre, 2013; Slavin, 2008; Slavin 2020). Through use of the external research evidence, a greater number of schools should be able to make more informed decisions, identify strategies and promising programmes that may be effective in their context and, ultimately, improve educational outcomes for pupils (e.g. Brown et al., 2017; Cooper et al., 2009; Godfrey, 2014; Nelson & O'Beirne, 2014). However, there is currently very little evidence to suggest that the uptake of more evidence-informed approaches has a positive impact on pupil outcomes (Gorard, 2020).

There are many common and interrelated challenges associated with embedding research evidence in schools. In 2003, Hemsley-Brown and Sharp (2003) reported results from a systematic review which found that there is a lack of time and capacity for teachers to engage with research; underdeveloped research literacy skills within the profession; negative attitudes and opinions of research evidence; and, a lack of accessible, relevant and usable evidence. Since this review, several other reviews and investigations have found that the same barriers continue to persist (see Cain, 2016; Gorard, 2020; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018; Walker et al., 2019).

If the external research evidence is to be used to improve teaching practice and provision then it is essential that there is high quality evidence available that is of practical use (Van Schaik et al., 2018; Walker et al., 2019). However, there is little agreement about what constitutes high-quality evidence in education. Many researchers place evidence derived from randomised controlled trials (RCTs) at the top of an evidence hierarchy, claiming it to be the gold standard of research (e.g., Goldacre, 2013; Slavin, 2008). However, others express concern about the overreliance of RCTs, which neglect to address the multiple and complex contextual issues associated with schools, and with the pupils and teachers within it (e.g. Hammersley, 2009, 2015; Thomas, 2016, 2021). Empirical evidence derived from experimental research can provide information about what programmes have the greatest chance of working ("what works"), but it is argued that teachers need more than this information to make sense of what will work in their context. Importantly, teachers also need to know if a programme or teaching approach will be effective for pupils in their school and, specifically, how they can get the approach to work in their setting to gain maximum impact with the available resources. A systematic review by Connelly et al. (2018) of all RCTs conducted in education from 1980 to 2016 also found that two thirds of RCTs neglect to address some of the pertinent factors such as implementation and causal process. It is, therefore, unsurprising that some teachers perceive some research evidence to be of little practical use (Van Schaik et al., 2018). Arguably, through greater use of mixed methods, researchers can determine the likely impact, causal process, the factors that impede or enhance implementation, and the acceptability and suitability of programmes and teaching approaches (Nutley et al., 2013).

If teachers are to use the external research evidence, then the evidence must be usable and trustworthy (Van Schaik et al., 2018). For example, despite limited research literacy skills, teachers should be able to draw conclusions from the results, apply this to the classroom, and have confidence in the findings. There is a proliferation of research evidence in education, and often popular programmes and teaching approaches have large bodies of evidence that is fragmented and difficult to access. Even when the evidence is available, research reports are often ambiguous and contradictory (Broekkamp & Van Hout-Wolters 2007). For many teachers, this makes evaluating the evidence challenging, time consuming, and often inconclusive. Without the necessary research literacy skills or time to fully understand and evaluate the validity, reliability, and integrity of the evidence, it becomes difficult to determine the likely impact of a programme. This may lead to school leaders – and policy makers more widely - choosing programmes that are based on weak causal evidence from methodologically flawed research (See, 2018). It is therefore necessary for researchers to invest time ensuring the research design and methods are transparent and limitations or threats to validity and reliability are clearly communicated (Bryman et al., 2006).

The ability to critically examine the design and execution of research is important for determining the trustworthiness of findings and understanding what counts as high-quality and low-quality evidence (Gorard et al., 2020). This is particularly relevant when synthesising or reviewing a body of evidence. According to Gorard et al. (2020) it is common for reviewers of research to neglect to attend to the quality and commensurability of studies, when reviewing or aggregating the results. This often leads to untrustworthy research findings and misleading conclusions drawn by educators who lack the research training in how to identify important methodological weaknesses in research outputs, even when presented in summary format via more trusted evidence repositories. There are many systems used for evaluating the quality of research. The Grading of Recommendations Assessment, Development, and Evaluation approach (GRADE) is one of the more popular systems, mandated by The Cochrane Collaboration (Movsisyan et al., 2018). The GRADE system rates the quality of evidence from high to low based on study design, limitations, inconsistency of results, indirectness of evidence, imprecision, and reporting of bias (Balshem et al., 2011). Gorard (2014) has also developed a system that evaluates the trustworthiness of research based on design, scale, attrition, quality of data or measurement quality, fidelity, validity and integrity. It is recommended that researchers now routinely employ evidence ‘sieves’ such as these to evaluate the quality of research findings, particularly when producing summaries of findings for teachers. The routine adoption of more robust assessment systems will ensure decision makers have access to meaningful and representative research outputs that will enable them to identify strategies and interventions that are more likely to have a positive causal impact on learner outcomes and, conversely, to avoid approaches that lack sufficient evidence of positive causal impact.

Over recent years important steps have been taken to identify the key facilitators that enable evidence uptake and mobilisation in schools (Cooper et al., 2009; Gorard et al., 2020; Nelson & Campbell, 2017; Schaik et al., 2018; Slavin, 2017), including: the availability of a range of high-quality and comprehensive evidence that evaluates a broad range of teaching approaches and programmes; the translation of evidence into accessible, teacher-friendly reviews; the need for greater cooperation and collaboration between researchers, teachers and policy makers; the need for schools to build organisational structures and foster cultures that support teacher enquiry and the use of research and evidence; and, the need for schools to be incentivised to adopt evidence-based programmes. Notably, findings from a recent review by Gorard et al. (2020), demonstrate that very little empirical evidence exists that supports the facilitators of evidence use. However, evidence suggests that although simply disseminating more accessible and practical summaries of evidence alone might improve users' knowledge, it does not seem to be an effective strategy for impacting practice.

There have been efforts from governments and researchers over recent years to improve the mobilisation of research evidence to improve outcomes for learners. A significant initiative by the United Kingdom government in 2010 was to commence funding 'what works' centres such as the Education Endowment Foundation (EEF) and the Early Intervention Foundation (EIF), intended to be trusted and accessible sources of what works evidence for school leaders and educators. An essential resource from the EEF is the Teaching and Learning Toolkit (https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/; Higgins et al., 2016). The aim of the toolkit is to provide schools with evidence-based summaries of the most effective teaching strategies and programmes, including information on set-up and implementation costs. Teachers can now also engage with educational research and find best practices methods through a variety of informal online groups and Twitter sites, and teacher-led initiatives such as 'ResearchED' (https://researched.org.uk/) and through professional organisations such as England’s Chartered College of Teaching (https://chartered.college/).

Through a range of other national and international educational research repositories such as the What Works Clearinghouse (https://ies.ed.gov/ncee/wwc/), and Evidence 4 Impact (https://www.evidence4impact.org.uk/), school leaders and teachers now have greater access to research evidence that collates summaries on a wide range of intervention programmes and approaches. However, much of this research has been conducted in primary schools, and there is a more limited evidence-base for educators in secondary schools (DfE, 2017). When Ager and Pyle (2013) sought the views of school leaders in England on their attitudes to accessing evidence for decision making, 67 per cent indicated they consult these types of evidence repositories when deciding how to spend their additional Pupil Premium funding, with nearly half referring to the EEF Toolkit. However, it has been suggested that efforts to engage with research may, in part, also be driven by a desire to satisfy accountability demands from inspection agencies (Brown & Greany, 2018). Brown and Greany reported more encouraging findings from teacher engagement with research evidence, with 76 per cent of teachers choosing to strongly agree or agree that research plays an important role in informing their practice, and 86 per cent indicating that information from research is useful in helping them apply new classroom approaches.

Despite a growing awareness of the importance of using research evidence to inform practice, and an increase in accessible information about ‘what works’, recent evidence suggests that evidence from external research is still infrequently used by teachers to inform teaching practice (Nelson et al, 2017). Findings from a series of teacher surveys has suggested that teachers are most likely to use past experiences of what works, and the experiences of colleagues within schools and from other schools, when choosing which teaching approaches to adopt (Nelson et al., 2017; Walker et al., 2019). These findings support previous studies, and suggest that teachers may be informed and influenced by trends, and the opinions and experiences of other teachers within their own and other schools (Greany & Brown, 2017). It would seem that when selecting interventions school leaders often favour programmes that are popular and compatible with past or current practice, despite the fact that these approaches often have little or no research evidence demonstrating that they have a causal impact on improving learner outcomes (Greany & Brown, 2017). Also, some evidence-based programmes may be viewed as incompatible with the school context in which they are to be implemented, and this prevents their adoption by teachers. For example, teachers may perceive interventions to conflict with a school's policies, philosophies and existing programmes or initiatives, and other approaches might be viewed as incompatible with available resources such as school budget, staff availability and available teaching materials (Bumbarger, & Perkins, 2008; Forman et al., 2009; Powers et al., 2010).

In addition to planning and delivering whole class instruction, teaching staff are routinely responsible for providing enhanced support to pupils with additional learning needs (ALN) in Wales, special educational needs (SEN) in England and Northern Ireland, and additional support needs (ASN) in Scotland. Pupils requiring this level of enhanced support comprise 22.2 per cent of all pupils in Wales, and 14.9 per cent of the school population in England (DfE, 2019; Welsh Government, 2019). In England, around 12 per cent of pupils with SEN are in primary schools and 10 per cent in secondary schools with a significant number of these pupils possessing a statement of Special Education Need or an Education, Health and Care plan (EHCP). Although enhanced provision for these pupils typically covers supplementary teaching for literacy and numeracy, the most common types of support are provided to pupils with: speech, language and communication difficulties; moderate or general learning difficulties; and, social, emotional and mental health difficulties (DfE, 2019;

Welsh Government, 2019). Moreover, there have been increased pressures to support pupils' mental health problems, which is reflected in recent government policies and guidelines (e.g. Green et al., 2004; Public Health England, 2017; Welsh Government, 2017).

Schools are now expected to provide effective support for pupils with additional needs and/or in need of supplementary teaching (e.g. DfE, 1997, 2016a, 2016b; Welsh Government, 2015b), and supplementary teaching is frequently delivered to small groups of pupils with similar needs and/or to pupils who require more intensive, individualised support (Estyn, 2020). This requirement is now closely linked with the provision of additional funding for schools to support disadvantaged and/or struggling learners (e.g. the Pupil Development Grant [PDG] in Wales and the Pupil Premium funding in England (DfE, 2015; Welsh Government, 2015b), and has led to the growth of the number and variety of interventions available to schools. Some interventions have robust causal evidence, many have preliminary evidence, some are ineffective, and a vast number simply have no evidence (see https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects; See & Gorard, 2020). As we have described previously, there is considerable value, and a moral imperative, to use the best available research evidence to identify strategies and interventions that are likely to have the greatest impact on pupil outcomes. This approach also ensures that scarce resources (including staff time and additional catch-up funding for disadvantaged learners) are spent more efficiently, and staff are not burdened with the additional and unnecessary task of implementing unproven or ineffective interventions (Gorard, 2020; Styles & Torgerson, 2018).

Very little information is available about the range and evidence supporting teaching interventions used in schools. Many published papers write about schools' and teachers' engagement with research evidence (e.g., Bell et al., 2010; Coldwell et al., 2017; Nelson et al., 2017; Walker et al., 2019), but few offer insights into the extent that research-evidence is in use in practice and, more specifically, how many interventions used in schools are evidence-based. To date, there has been no systematic attempt to better understand the range and evidence base for interventions that are being used in UK schools.

The following study was conducted with a well-established cluster of schools in Wales to evaluate the range and evidence base for the interventions used across the two secondary schools, seven primary schools and one special school. In the interest of moving towards a more evidence-informed practice, the cluster of schools had entered into a research partnership with the Regional School Effectiveness and Improvement Service for North Wales (GwE) and the Collaborative Institute for Education Research, Evidence and Impact (CIEREI), Bangor University, to commission this evaluation work alongside other close-to-practice research projects. Throughout this study, the head teachers and senior leaders were closely involved in the development and planning of the research and worked collaboratively with the lead researcher who was embedded within the cluster.

The primary aim of this study was to explore the range of interventions used in the cluster and to assess the evidence base for the interventions, in particular to help identify the interventions and/or programmes that possess evidence of positive causal impact on pupil outcomes. A secondary aim was to investigate whether increased knowledge about the evidence for the interventions used in the cluster, presented through a summary report to school leaders, had an impact on existing provision. The intention of this collaborative research project was to help move the cluster towards a more evidence-informed approach whereby school leaders will be better able to make more informed decisions on school provision by drawing on a more trustworthy summary of the available research evidence. These findings will also provide policy makers with an improved understanding of the range and evidence base of provision in a school cluster and will inform discussions around how best to facilitate evidence-use in schools. The aim of this study was to answer the following questions:

1. What is the range of interventions used across a school cluster?
2. What is the evidence-base for these interventions?
3. How will access to this information impact provision in schools?

Method

Design and analysis

In order to get a robust answer to the research questions asked we used mixed methods design incorporating questionnaire, systematic review and qualitative interviews. First, to understand the range of interventions used across a school cluster, data were collected using questionnaires, then categorised and analysed using descriptive statistics. Second, to assess the causal evidence for the interventions, a systematic review of the literature was conducted. Finally, for the follow-up investigation, checklists of interventions were used and interviews were conducted to identify any changes that were made to school provision. Questionnaire data was analysed using descriptive statistics, and the interview data was analysed using thematic analysis, as outlined by Braun and Clarke (2006). For the thematic analysis, the recorded data was transcribed and read twice before, salient and interesting comments were coded and converted into themes. The final and main themes are presented in a narrative description.

Ethics

The study was conducted under ethical approval from the School of Education Research Ethics Committee, at Bangor University (18-03). We obtained consent from each participating head teacher. To protect anonymity, identifiable information such as school or head teacher names are not disclosed.

Participants

The participants in this study were a cluster of 10 schools located in an area of high deprivation in an urban area of the United Kingdom. The cluster consists of seven primary schools (six English medium and one Welsh medium), two secondary schools (English medium), and a special needs school (English medium), and represent a typical group of medium to large schools in a large town. For each school, the head teacher and a member from the senior leadership team, considered to be most knowledgeable about school provision, were responsible for completing the survey of interventions.

Follow up investigation

Nine schools took part in the follow-up investigation to assess the impact of the information about the evidence for interventions on school provision. One school withdrew because of a change in leadership. The head teacher and a member from the senior leadership team involved in the primary data collection were responsible for completing it at follow-up.

Data collection methods

We developed a questionnaire to collect information about the interventions currently in use in each school. Because so little is currently known about the range of interventions used in schools, this study intentionally kept the inclusion criteria broad. This decision was guided by a requirement to identify all relevant interventions, regardless of type. The questionnaire included a series of demographic questions, and a question that asked respondents to provide a list of all school-based interventions used in the current academic year. A spreadsheet was provided for participants to input the interventions identified. Alongside each intervention participants were asked to indicate how the intervention was used in the school, e.g. universal provision (tier 1), group-based support (tier 2) and/or one-to-one intensive support (tier 3).

Follow-up investigation

To assess the impact of the information about the evidence for interventions on school provision, a checklist of interventions was created for each school. This was a list of all the interventions that were primarily collected from the school. The respondent was asked to place a mark next to the intervention to indicate if the intervention was ‘still in use’ or ‘not in use’. In addition, the head teachers were invited to attend a short one-to-one interview with the first author.

Procedure

We collected data using a multiphase iterative approach that included three phases. This approach allowed the researcher to engage with each data collection phase in a reflexive way, repeating steps to ensure the final data set was comprehensive, and that any newly implemented interventions were not missed. A copy of the questionnaire and spreadsheet were emailed to the head teachers in February 2018. The data were collected across three time points; phase one (May 2018), phase two (June 2018) and phase three (July 2018). To encourage responses, schools were sent reminder emails before each phase. Throughout the process the researcher met with the head teachers in their termly cluster meetings to support the data collection process and answer any questions on the inclusion criteria for an intervention.

In phase one respondents were asked to provide a list of all school-based interventions. In phase two, the research question was redefined to remove the ambiguity of the term ‘intervention’ and to remove any limitations that might have prevented an intervention from being included. Teachers were instead asked to provide an up-to-date list of all school-based ‘provision’. In phase three, schools were again asked to provide an updated list of school provision, and in addition, were offered an un-structured meeting with the researcher, to facilitate the process. Three schools participated in a one-to-one meeting with the researcher, in which questions were asked about the practices and programmes used with pupils outside of routine teaching. The final list of interventions was collated at the beginning of July 2018.

In October 2018, schools were provided with simple individualised summary reports. Each report included a table of interventions used by the school and provided a summary of the evidence base that was gathered following a systematic review of available literature (e.g. reports summarised whether there was preliminary, promising or no evidence of causal impact for each intervention).

Follow-up investigation

To assess the impact of the information about the evidence for interventions on school provision we sent a checklist of interventions to each of the head teachers at the end of November 2019. Head teachers were offered the choice of completing and returning the follow-up checklist by email (n = 2) or completing with the first author during a one-to-one interview (n = 2). Head teachers also received an invite to attend a one-to-one, semi-structured interview with the first author to answer the following question: has greater information about the evidence-base of interventions had an effect on the way you choose school provision?

Screening of the interventions

The total number of interventions collected from the cluster was 242. This included 157 unique interventions and 85 duplicate interventions, across all 10 schools (e.g. Emotional Literacy Support Assistant [ELSA] was a duplicate intervention used in six of the schools).

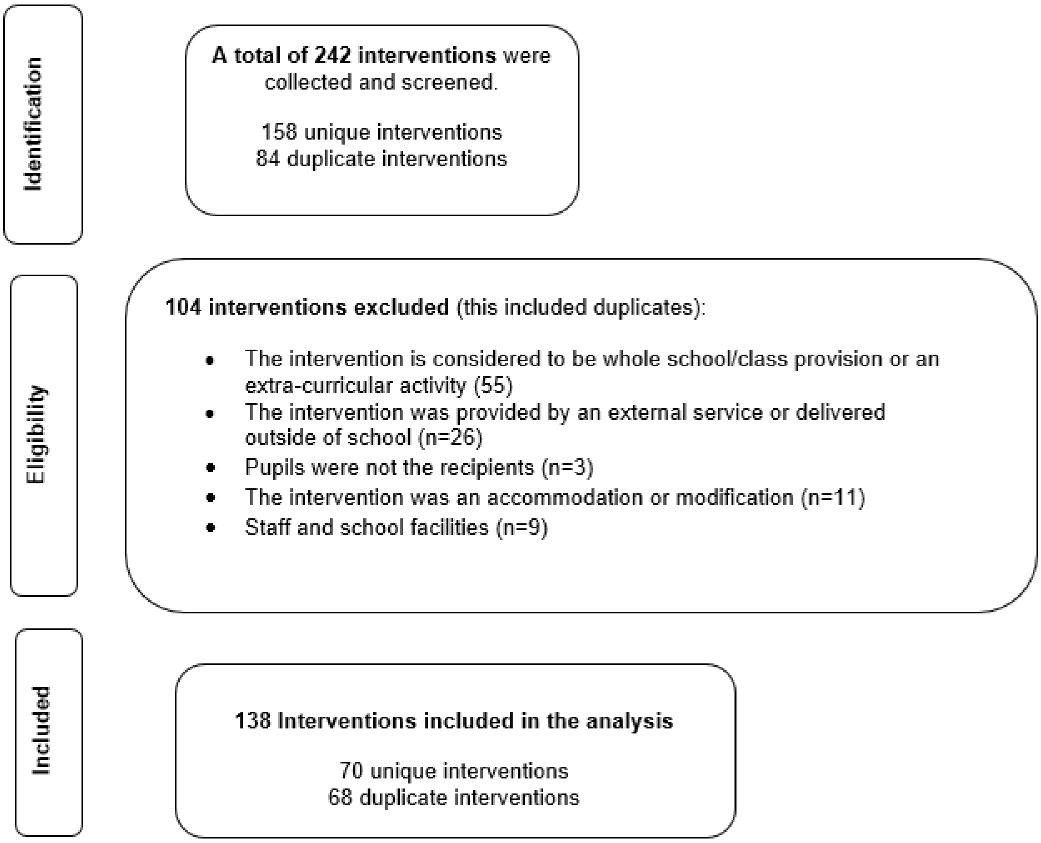
The interventions used by the schools were initially screened for inclusion in the study according to the inclusion and exclusion criteria, and those that did not meet the criteria for an intervention were removed from the study. These judgements were also independently assessed by another researcher, any discrepancies were discussed, and agreement was reached on 100 per cent of the data.

Definition of a school-based intervention

We could not find a widely accepted single definition for the term ‘intervention’ in education research nor in the wider school system. For this study we have used a definition derived from the Response to Intervention (RtI) tiered approach (Fletcher & Vaughn, 2009). A school-based intervention is defined as the provision of supplementary support, targeted at addressing deficits in pupil learning and development, when whole class instruction or the typical universal teaching provision is not sufficient (described in the RtI model as tier 1 or standard provision). Interventions are different from, or additional to the universal standard provision offered by the school or class teacher to all children and are typically delivered to small groups of pupils with similar need (described as tier two interventions), or individual pupils with persistent and significant need (described as tier three interventions). Interventions are not accommodations (e.g. a change to the classroom environment that gives students equal access to learning), or modifications (changes to what a student is taught or expected to learn).

In accordance with the definition of a school-based intervention, interventions were included if: they were delivered in school, during school hours; pupils aged 4–18 years were the recipients; and, they were used as small group (tier two) and/or one-to one (tier three) supplementary support that targeted deficits in learning, well-being and/or behaviour.

Interventions were excluded if: they were used as universal/whole school provision (tier one); pupils were not the recipients; they were considered to be an accommodation or modification; they represented teaching staff or school facilities; and/or they were an external service. Figure 1 shows the results from the screening and eligibility process and Table 1 shows examples of interventions that were removed from the study. For a full list of the interventions excluded see Appendix A.



A total of 242 interventions were collected and screened.

158 unique interventions
84 duplicate interventions

104 interventions excluded (this included duplicates):

- The intervention is considered to be whole school/class provision or an extra-curricular activity (55)
- The intervention was provided by an external service or delivered outside of school (n=26)
- Pupils were not the recipients (n=3)
- The intervention was an accommodation or modification (n=11)
- Staff and school facilities (n=9)

138 Interventions included in the analysis

70 unique interventions
68 duplicate interventions

Figure 1. Flow diagram displaying the number of interventions included and excluded from the study.

Table 1 *Examples of interventions excluded from the study, following screening*

Provision excluded	Reason for exclusion
Personalised timetable	Accommodation or modification
Sensory breaks	Accommodation or modification
Visual Timetable	Accommodation or modification
Reduced hours	Accommodation or modification
Behaviour support services	External
Child and Adolescent Mental Health Services	External
Counselling service	External
Educational Psychologist	External
Key worker	Teaching staff
Safe-Guarding Team	Teaching staff

Higher Level Teaching Assistant	Teaching staff
Learning Coach	Teaching staff
Young Carers in Schools Programme	Pupils were not the recipient
Direct Instruction	Universal provision
After-school clubs	Universal provision
Social and Emotional Aspect of Learning (SEAL)	Universal provision
Attendance rewards	Universal provision

Categorising the interventions

A total of 138 interventions were included in the final analysis (see Figure 1 and Table 2). This included 70 unique interventions and 68 duplicate interventions (e.g., some of the unique interventions were used by more than one school). In response to the broad range of interventions collected, we decided to adopt the ‘areas of need’, as defined in the SEN Code of Practice, as a means of broadly categorising the interventions. These areas of need are generally well-understood by educators and policy makers throughout the United Kingdom, and provide a useful framework for categorisation. The areas of need are: Communication and Interaction; Cognition and Learning; Social, Emotional and Mental Health; and Sensory/or Physical (DfE, 2015).

First, we collated interventions into an Excel spreadsheet, then coded and categorised them according to the SEN areas of need. Next, we categorised the interventions into sub-categories according to their instructional focus, which was determined by searching programme websites. The term ‘non-specific’ was adopted for interventions that did not name a specific programme or include a defined instructional approach (e.g. nurture groups consist of groups of vulnerable learners created without the use of specific protocols or approaches).

Rapid systematic review of the evidence

We conducted a rapid systematic review of the literature to assess the evidence-base for the interventions. Any interventions defined as non-specific were excluded from the systematic review because research could not be accessed for unnamed programmes and approaches. We could therefore make no comment on the effectiveness of these programmes.

Search strategy

We completed the first search of the literature in 2018. To ensure new studies were included, a second search was completed in 2020. The search strategy relied on three main sources.

First, we completed searches using the following electronic search engines and their included sociological, psychological and educational databases, using the intervention name (e.g. “Headsprout”):

Web of Science – Web of Science Core Collection, Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, Emerging Sources Citation Index, Conference Proceedings Citation Index, Book Citation Index.

EBSCOhost - CINHAL, Library, Information Science & Technology Abstracts and MEDLINE).

ProQuest - PsycINFO, ProQuest Dissertations & Theses Global, Social Science Premium Collection, Education Collection and ERIC, and Social Science Database.

Based on the preliminary analysis of the availability of the literature, key words were added to searches to narrow or broaden the results. In addition to primary studies, we also obtained reviews from the searches, and searched reference lists, and if relevant, references were retrieved.

Second, we conducted searches in a range of national and international research websites using the name of the intervention as the search term (e.g. “Seasons for Growth”):

These were the Early Intervention Foundation, Education Endowment Foundation; The What Works Clearinghouse, administered by The Institute of Educational Sciences (IES) of the U.S. Department of Education; Best Evidence Encyclopaedia, National Registry of Evidence-Based Programs and Practice; The Collaborative for Academic, Social, and Emotional Learning (CASEL); Evidence4impact; The Communication Trust, National Foundation for Educational Research and Blueprints for Violence Prevention Model and Promising Programs, administered by the Center for the Study and Prevention of Violence at the University of Colorado; and, The National Registry of Evidence-Based Programs and Practices, administered by the Substance Abuse and Mental Health Services Administration (SAMHSA), a branch of the U.S. Department of Health and Human Services. Reviews were obtained from the websites, and, where relevant, citations retrieved.

Finally, if very little or no evidence was found for a named intervention, then we searched Google Scholar and programme websites.

Literature selection

In this review, we aimed to identify the causal evidence for each named intervention programme or instructional approach. The review drew upon national and international published research and evidence, all of which were reported or published in English from 2001 onwards. Primarily, research evidence was included if the intervention had a clearly identifiable name (e.g. “Catch-up Literacy”), the research was based on pupil outcomes (4-18 years) and was undertaken in the school. The studies of particular interest in this review were those that sought to establish a cause-effect relationship and, therefore, only experimental randomised control trials and quasi-experimental studies (non-equivalent groups and no comparison group [pre/post]) were included. Studies were excluded if they were non-empirical, case studies, if the intervention in question was combined with another intervention, not based on pupil outcomes and conducted outside the classroom.

Screening and data extraction

A total of 2596 studies were retrieved following searches. The studies were first screened for relevance by examining the title and abstract and applying the inclusion and exclusion criteria. Once irrelevant and duplicate studies were removed, a total of 168 remained. Studies that were retained were retrieved and the full text was read and assessed for eligibility. The 43 studies that did not meet the inclusion criteria were excluded, leaving 125 included studies. The following data was then extracted from the retained studies: the research design, population, sample size, outcomes, attrition levels, outcome measures used, and any methodological limitations that might threaten validity (e.g. non-equivalent groups, missing data, unsuitable analysis methods and conflicts of interest). Reasons for exclusion as well as the selection process are displayed in Figure 2.

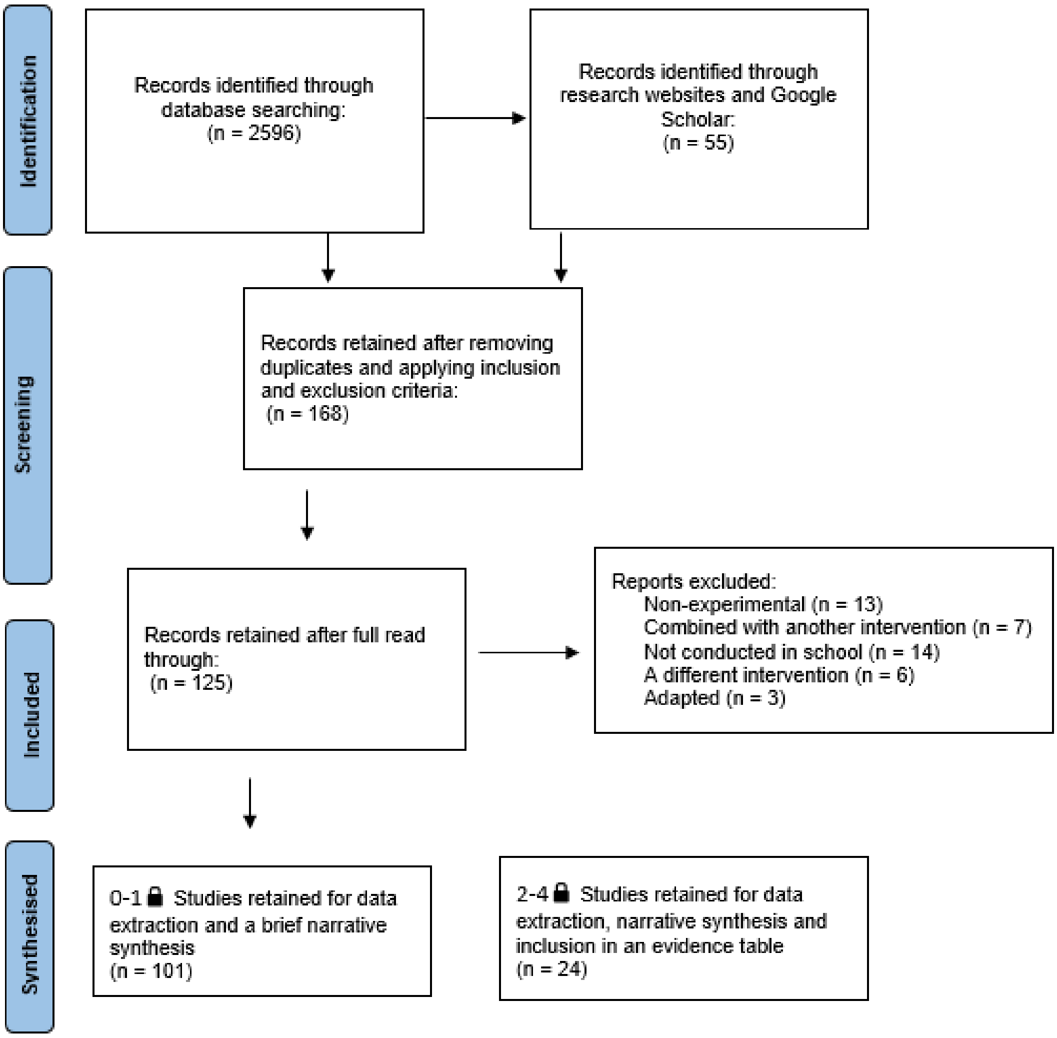

Identification
Records identified through database searching: (n = 2596)
Records identified through research websites and Google Scholar: (n = 55)
Screening
Records retained after removing duplicates and applying inclusion and exclusion criteria: (n = 168)
Included
Records retained after full read through: (n = 125)
Reports excluded:
Non-experimental (n = 13)
Combined with another intervention (n = 7)
Not conducted in school (n = 14)
A different intervention (n = 6)
Adapted (n = 3)
Synthesised
0-1 🔒 Studies retained for data extraction and a brief narrative synthesis (n = 101)
2-4 🔒 Studies retained for data extraction, narrative synthesis and inclusion in an evidence table (n = 24)

Figure 2 Flow diagram illustrating study selection process (adapted from PRISMA diagram—Moher et al., 2009)

Quality assessment

To judge the quality of group design research, we used the ‘sieve’ developed by (Gorard et al., 2017). Each study was given an individual rating from 0🔒, which represents a study that adds nothing to our knowledge of social science; to 4🔒, which represents a robust study. The rating is based on five criteria: study design, scale of study, bias through missing data, quality of data obtained and other threats to validity.

Evidence synthesis and presentation of the results

The results section below contains a narrative of included studies for each intervention, and a summary of the evidence retrieved. As a result of the breadth of evidence identified, we only briefly summarised single-case design studies, and group design research rated 0-1🔒, in the narrative. The remaining evidence, judged to be of moderate and higher quality according to the ‘sieve’ ratings (2🔒 and above), we evaluated more extensively in the narrative and presented the extracted data in an accompanying evidence table, along with study quality ratings.

Results

The following section will report on the range of interventions used across the cluster.

A total of 138 interventions were in use across the 10 schools. See Table 2 for the frequency of interventions in each school organised by SEN category. Across the cluster, the percentage of pupils reported as having Additional Learning Needs (ALN) was 34 per cent.

Table 2. *Frequency of interventions in each school organised by SEN 'Areas of Need'*

School	No. pupils	No. of pupils with ALN (SEN)	All	Cognition & learning	Social, emotional & mental health	Communication & interaction
Primary						
1	702	148	25	12	7	**6
2	480	67	15	8	5	2
3	470	297	16	7	9	0
4	446	191	24	7	11	**6
5	308	68	9	5	3	1
6	183	45	11	5	5	1
*7	478	106	4	-	4	-
Secondary						
1	260	85	15	7	7	1
2	980	379	9	1	8	-
Special needs	137	137	10	3	5	2
Total	4444	1523	138	55	64	19

* Welsh medium school, **speech and language services at the school

Cognition and Learning

There were 55 cognition and learning interventions in use across the cluster. Across the 10 schools, 38 interventions focused on literacy. Nine schools reported offering one or more literacy intervention. The median number of literacy interventions was 4 (IQR = 2 - 5). The most common type of literacy programme focused on reading instruction (17), followed by multi-skill literacy programmes, which incorporate reading, writing, and oral language instruction in combination (11). The remaining literacy programmes focused on phonics (3), spelling (1) and handwriting (1). Five interventions were defined as 'non-specific', and these were used to teach handwriting (1), spelling (1), reading through high frequency word acquisition (1), and general literacy catch-up (2). The most commonly used literacy programmes were Headsprout (used in five schools), Read Write Inc. phonics (used in four schools) and Catch-Up Literacy (used in four schools).

There were 17 numeracy interventions in the cognition and learning category. Eight schools offered one or more numeracy interventions. The median number of interventions was 2 (IQR = 1 - 2). Fourteen numeracy interventions focussed on developing a variety of mathematical skills (multi-skill) and one was devoted to multiplication practice. The remaining two numeracy interventions were ‘non-specific’ and these were used for numeracy catch-up sessions. The most popular numeracy intervention was Say All Fast a Minute Every Day (SAFMEDS), a precision teaching strategy, used by eight schools to improve fluency in basic number skills. Table 3 shows the list of cognition and learning interventions used across the cluster.

Table 3. *Cognition and learning interventions used across the cluster*

Intervention	Frequency of interventions used in the cluster	Instructional focus
Numeracy		
SAFMEDs (manualised procedure)	8	Multi-skill
Sumdog	2	Multi-skill
RM Easi maths	1	Multi-skill
YIPIYAP Numeracy	1	Multi-skill
Number shark	1	Multi-skill
My Maths	1	Multi-skill
Times Table Rockstars	1	Multiplication
Non-specific		
Numeracy booster group	2	
Literacy		
Accelerated reader	1	Reading
Catch Up Literacy	4	Reading
Headsprout	5	Reading
Reading Educational Assistance Dogs	1	Reading
Reading Rocks	1	Reading

Toe by Toe	3	Reading
Wolf hill	1	Reading
IDL Literacy	1	Multi-skill
Nessy - Dyslexia	1	Multi-skill
Read Write Inc. Phonics	4	Multi-skill
Read Write Inc. Fresh Start	1	Multi-skill
Word Shark	2	Multi-skill
Word wasp	1	Multi-skill
YIPIYAP Literacy	1	Multi-skill
Bugs Reading intervention	1	Reading
Direct Phonics - DP Publishing	2	Phonics
Silly Sounds	1	Phonics
Hand to Spell	1	Spelling
Write from the start, Teodorescu	1	Handwriting
Non-specific		
Targeted handwriting	1	
Precision & high frequency words	1	
Targeted spelling	1	
Literacy booster group	2	
Total no. of interventions	55	

Social, emotional and mental health

There were 64 social, emotional and mental health interventions used across the cluster. Ten schools offered one or more interventions in this category. The median number of interventions was 6 (IQR = 5–8). Thirty-four interventions were programme specific (e.g., Incredible Years, Dinosaur school) and included two therapeutic support programmes, two health and well-being programmes, an approach to reduce problem behaviour, and 30 programmes that focused on developing pupils’ social and emotional skills. Nine schools offered one or more programmes that focused on developing pupils’ social and emotional skills. The median number of interventions was 3 (IQR = 2–4.50). The most common programmes used to teach social and emotional skills were Emotional Literacy Support

Assistants (ELSA) (used in six schools), followed by Seasons for Growth (used in five schools), and Friends for Life and Forest School (both used in four schools). Twenty-nine interventions were categorised as 'non-specific' (e.g., anger management). All 10 schools offered interventions in the social, emotional and mental health category that were non-specific. The median number of interventions was 3 (IQR = 2–4). The most common non-specific intervention was the nurture group, and there were 10 nurture groups implemented across seven schools. See Table 4 for the list of interventions in this category, used across the cluster.

Table 4. *Social, emotional & mental health interventions used across the cluster*

Intervention	Frequency of interventions used in the cluster	Focus
Incredible Years, Dinosaur School	2	Social & emotional skills
Emotional Literacy Support Assistant (ELSA)	6	Social & emotional skills
Forest school	4	Social & emotional skills
Friends for Life	4	Social & emotional skills
Mindfulness in Schools project (MISP), Paws b	1	Social & emotional skills
Rhythm of life	1	Social & emotional skills
Seasons for Growth	5	Social & emotional skills
Talkabout	3	Social & emotional skills
Unearthing box	2	Social & emotional skills
Be Smart, Be Cool, Be Positive	1	Social & emotional skills
Squash the change	1	Social & emotional skills
Time-out	1	Improve behaviour
Drawing and Talking therapy	1	Therapeutic support
Colour Away Your Worries	1	Therapeutic support
Student Assistance programme (SAP)	2	Wellbeing support
Non-specific		
Five strikes contract	1	
Coastal school	1	
Mentoring	2	
Peer support	2	

Nurture groups	10
Anger management	4
Anxiety support	1
Bereavement	1
Emotional support	1
Growth Mind-set	1
Relaxation	1
Self-esteem support	2
Animal assisted therapy (dog)	1
Play therapy	1
Total no. of interventions	64

‘Non-specific’ is an intervention that is a non-specific programme or an undefined approach

Communication and interaction

Across the cluster, 19 interventions were categorised as communication and interaction provision. The intervention programmes focused on pupils’ speech and language skills (13), communication skills (1), socialisation and communication skills (1) and listening and attention skills (1). Seven schools offered one or more intervention that focused on communication and interaction difficulties. The remaining three interventions were non-specific. The median number of interventions was 1 (IQR = 0 – 2). The most commonly used speech and language interventions were Elklan and Picture Exchange Communication system (PECS) (both used in four schools), and WellComm (used in three schools). Table 5 shows the list of communication and interaction interventions used across the cluster.

Table 5. *Communication and interaction interventions used across the cluster*

Intervention	Frequency of interventions used in the cluster	Instructional focus
Language for Thinking	1	Speech & language
Social Stories	1	Speech & language
Wellcomm	3	Speech & language
Narrative Therapy Beck Shanks	4	Speech & language

Elklan, Talking Matters	4	Speech & language
Picture Exchange Communication System (PECS)	1	Communication
Lego based therapy - Bricks for Autism	1	Social skills & communication
Lola Speaking and Listening	1	Listening & attention
Non-specific		
Speech & Language therapy	2	
Social communication group	1	
Total no. of interventions	19	

‘Non-specific’ is an intervention that is a non-specific programme or an undefined approach.

The following section will report on the research evidence for the named intervention programmes or approaches used by the cluster. The descriptive statistics reported are based on the multiple occurrences of interventions used across the cluster (e.g., ELSA was reported as being used in six schools).

Results from the review of the evidence found that 30 per cent of all interventions in use across the cluster had causal evidence of impact on a pupil outcome; 67 percent had no evidence; and three per cent had evidence of ineffectiveness. Table 6 displays the descriptive statistics for the causal evidence of the interventions used across the cluster.

Table 6. *Summary of the causal evidence for interventions used across the cluster*

	Evidence of causal effect	Promising (2-4🔒)	Preliminary (0-1🔒)	No evidence	Ineffective
	Frequency (%)				
All interventions	42 (30)	15 (11)	27 (19)	92 (67)	4 (3)
Cognition & learning	19 (35)	10 (18)	9 (17)	32 (58)	4 (7)
Social, emotional & mental health	15 (23)	4 (6)	11 (17)	49 (77)	-
Communication & interaction	8 (42)	1 (5)	7 (37)	11 (58)	-

Five of the commonly used interventions had causal evidence of effect. Headsprout, Toe by Toe, and Narrative Therapy by Beck Shanks each had preliminary evidence, rated low quality (1🔒). SAFMEDS and Friends for Life had promising evidence, rated moderate to high quality (2-4🔒). The remaining programmes, which included Read Write Inc. Phonics, ELSA, Forest School, and Seasons for Growth had no evidence available at the time of review; and, Catch-Up Literacy had 3 🔒 evidence that found it to be ineffective (see Tables 7 - 9).

Table 7. *Summary of the causal* evidence *for the cognition and learning interventions*

Intervention	Frequency of interventions used in the cluster	Causal evidence	Quality rating
Numeracy			
SAFMEDs (manualised procedure)	8	Promising	2🔒
Sumdog	2	No evidence	
RM Easi maths	1	No evidence	
YIPIYAP Numeracy	1	No evidence	
Number shark	1	No evidence	
My Maths	1	No evidence	
Times Table Rockstars	1	No evidence	
Non-specific			
Numeracy booster group	2	No evidence	
Literacy			
Accelerated reader	1	Promising	2-3🔒
Catch Up Literacy	4	Ineffective	
Headsprout	5	Preliminary	0-1🔒
Reading Educational Assistance Dogs	1	Preliminary	0-1🔒
Reading Rocks program	1	No evidence	
Toe by Toe	3	Preliminary	0-1🔒
Wolf hill	1	No evidence	
IDL Literacy	1	No evidence	
Nessy - Dyslexia	1	No evidence	
Read Write Inc. Phonics	4	No evidence	
Read Write Inc. Fresh Start	1	Promising	3🔒

Word Shark	2	No evidence
Word wasp	1	No evidence
YIPIYAP Literacy	1	No evidence
Bugs Reading	1	No evidence
Direct Phonics - DP Publishing	2	No evidence
Silly Sounds	1	No evidence
Hand to Spell	1	No evidence
Write from the start, Teodorescu	1	No evidence
Non-specific		
Targeted handwriting	1	No evidence
Precision & high frequency words	1	No evidence
Targeted spelling	1	No evidence
Literacy booster group	2	No evidence

Table 8. *Summary of the causal evidence for the social, emotional and mental health interventions*

Intervention	Frequency of interventions used in the cluster	Causal evidence	Outcome
Incredible Years, Dinosaur School	2	Preliminary	0-1 🔒
Emotional Literacy Support Assistant (ELSA)	6	No evidence	
Forest school	4	No evidence	
Friends for Life	4	Promising	2-4 🔒
Mindfulness in Schools project, Paws b	1	Preliminary	0-1 🔒
Rhythm of life	1	No evidence	
Seasons for Growth	5	Preliminary	0-1 🔒
Talkabout	3	No evidence	

Unearthing box	2	No evidence	
Be Smart, Be Cool, Be Positive	1	No evidence	
Squash the change	1	No evidence	
Time-out	1	Preliminary	0-1 🔒
Drawing and Talking therapy	1	No evidence	
Colour Away Your Worries	1	No evidence	
Student Assistance programme (SAP)	2	Preliminary	0-1 🔒
Non-specific			
Five strikes contract	1	No evidence	
Coastal school	1	No evidence	
Mentoring	2	No evidence	
Peer support	2	No evidence	
Nurture groups	10	No evidence	
Anger management	4	No evidence	
Anxiety support	1	No evidence	
Bereavement	1	No evidence	
Emotional support	1	No evidence	
Growth mindset	1	No evidence	
Relaxation	1	No evidence	
Self-esteem support	2	No evidence	
Animal assisted therapy (dog)	1	No evidence	
Play therapy	1	No evidence	

Table 9. *Summary of the causal evidence for the communication and interaction interventions*

Intervention	Frequency of interventions used in the cluster	Causal evidence	Outcome
Language for Thinking	1	Preliminary	0-1🔒
Social Stories	1	Preliminary	0-1🔒
Wellcomm	3	No evidence	
Narrative Therapy, Beck Shanks	4	Preliminary	0-1🔒
Elklan, Talking Matters	4	No evidence	
Picture Exchange Communication System	1	Promising	2🔒
Lego based therapy - Bricks for Autism	1	Preliminary	0-1🔒
Lola Speaking and Listening	1	No evidence	
Non-specific			
Speech & Language therapy	2	No evidence	
Social communication group	1	No evidence	

Interventions with promising evidence

The following sections will summarise the evidence from the review and will focus on the intervention programmes and instructional approaches that have evidence rated as moderate quality and higher (2 and above🔒) that show more promising causal impact on pupil outcomes (see Tables 7 - 9).

Say All fast Minute Every Day Shuffled (SAFMEDS)

Say-All-Fast-Minute-Every-Day-Shuffled (SAFMEDS) is a precision teaching strategy that is often used for instruction and assessment (Quigley et al., 2018). This review identified a total of seven studies that demonstrated positive causal impact on pupil outcomes. Four studies used a single-case design. Two studies used a single-subject design with multiple baseline to measure maths fluency with learners in elementary school, with learning difficulties (Casey et al., 2003; Cunningham et al., 2012). Two studies used a single-subject multiple baseline design, of which one measured reading fluency, in a primary school, with typically developing learners (Lambe et al., 2015) and the other language acquisition, in elementary school, with learners that had emotional and behavioural problems (Shue, 2017). One very small study employed a single subject-design and randomly assigned students to conditions, to measure the reading fluency of secondary pupils (Hughes et al., 2007), and this was rated 0-1🔒. Another small-scale (n = 48) quasi-experimental study with non-equivalent groups measured maths fluency with secondary school pupils (Hunter et al., 2016), and this was rated 0-1🔒.

There was one quasi-experimental study, considered to be moderate quality (2🔒) that found large positive effects on language acquisition (Beverley et al., 2016). Although Beverley et al. (2016) had a medium number of participants, the study is considered weak in design because of the non-equivalent and imbalanced groups (see Table 10).

Each of the studies adds to a causal evidence-base. However, collectively, most of the studies were non-commensurable (measuring different populations and outcomes). Based on the current evidence, the SAFMEDS instructional approach is considered promising. We would recommend that research should focus on developing more robust and replicated randomised research to strengthen the promising evidence.

Table 10. *A summary of the evidence for SAFMEDS*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Beverley et al. (2016)	Quasi - experimental	Grade 7	*n*intervention = 79 *n*control = 16	Effective (1.88)	2🔒

Accelerated Reader (Renaissance Learning)

Accelerated Reader (AR) is a computerised reading programme that monitors and manages pupils' reading practices and fosters independent reading. There is a considerable evidence-base for this intervention (What Works Clearinghouse, 2007a, 2007b, 2009a, 2009b, 2010a; 2010b;) and this review included 36 studies, although 32 of these are judged to have weaknesses in design that prevents claims of causal effect (0-1🔒). Commonly, the studies used a causal comparative design (ex post facto) with non-equivalent groups. Other studies had no comparison group and measured impact using the Standardized Test for Assessment of Reading (STAR) that is produced and marketed by Renaissance Learning and embedded in the AR programme. Some of this research finds positive effects on pupil outcomes (e.g., McGlinn & Parrish, 2002; Nunnery et al., 2006; Nunnery & Ross, 2007; Rodriguez, 2007; Topping & Fisher, 2003). Notably, some studies show no impact from using AR, and a number report greater gains in the control groups (Huang & SuHua, 2011; Melton et al., 2004; Pavonetti et al., 2002).

The current study considered four studies to be moderate quality (2-3🔒) and with mixed findings (see Table 11). In a small-scale RCT by Bullock (2004), results from reading fluency and reading comprehension tests using the STAR and Dynamic Indicators of Basic Early Literacy Skills (DIEBELS) found no significant effects. The large quasi-experimental study by Ross et al. (2004) that used only the STAR literacy test also found no effects for students in grades 4-6 (age 9 to 12 years) but did find moderate and large positive effects for pupils in grades K-3 (age 5 to 9 years). Ross et al. (2004) did establish baseline equivalence, but also reported very large attrition rates. Shannon et al., (2015) found significant moderate effects on reading and, although groups were non-equivalent, baseline equivalence was reported. However, the results were obtained using only the STAR assessment tool. The remaining RCT demonstrated small positive effects from using AR on reading ability using The New Group Reading Test (NGRT), and negligible attrition was reported (Siddiqui et al., 2016).

Based on the more robust evidence (2-3🔒) available for AR, it is considered to be a promising intervention.

Table 11. *A summary of the evidence for Accelerated Reader*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Bullock (2005)	RCT	Grade 3-5	*n*intervention = 55 *n*control = 59	No effect	2🔒
Ross et al. (2004)	Quasi-experimental	Grade k-5	*n*intervention = 615 *n*control = 442	Effective for grade k-3 (0.71, 0.36, 0.25, 0.33); non-significant for grade 4-6 (0.01, 0.11, 0.14)	2🔒
Shannon et al. (2015)	Quasi-experimental	Grade 1-4	*n*intervention = 189 *n*control = 157	Effective (0.38)	2🔒
Siddiqui et al. (2016)	RCT	Year 7	*n*intervention = 166 *n*control = 183	Effective (0.24)	3🔒

Read Write Inc. Fresh Start

Read Write Inc. Fresh Start is a catch-up literacy programme that uses a systematic synthetic phonic approach to teach reading skills to learners at the end of primary and the beginning of secondary school. This review found one moderately robust RCT based on Fresh Start (Gorard et al., 2016; 3🔒). The school-led study reported small positive gains in reading comprehension. However, the study also reported some imbalance in the groups, which prevented a higher quality trustworthiness rating (see table 12). Although Read Write Inc. is considered to be a promising programme, we would recommend future research is focused on replication of the research to strengthen causal claims.

Table 12. *A summary of the evidence for Read Write Inc. Fresh Start*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Gorard et al. (2016)	RCT	Year 7	*n*intervention = 212 *n*control = 221	Effective (0.24)	3🔒

Friends for Life (FRIENDS)

Friends for Life is a mental health program that can be used to teach emotional resilience and/or to address child and adolescent anxiety and depression. Friends has a substantial international evidence-base and has been recognised by the World Health Organization as an effective tool (see the following reviews: Briesch et al., 2010; Higgins & O'Sullivan, 2015; Maggin & Johnson, 2014). This review included 26 studies, of which 11 were rated low quality (1🔒) due to weaknesses in the research design. For example, one single-case multiple baseline design study reported positive changes for pupils with emotional, behavioural and/or anxiety disorders (Schoenfeld & Mathur, 2009). Two further small-scale studies used a pre/post design with no control condition and found positive effects of FRIENDS when used selectively (Cooley et al., 2004; Fjermestad et al., 2020). Another moderate-scale pre-post design and its follow-up study both reported positive effects from universal delivery of FRIENDS (Stallard et al., 2005, 2008). Conversely, two small-scale quasi-experimental studies found no statistically significant effects post intervention, either when used selectively (Mostert & Loxton, 2008) or universally (Rose et al., 2009), although Mostert and Loxton (2008) did report positive effects at follow-up. Finally, four very small-scale RCTs found significant positive effects on measures of anxiety when used universally (Rodgers & Dunsmuir, 2015) and selectively (Bernstein et al., 2005; Liddle & Macmillan, 2010; Siu, 2007).

The current study considered 11 studies to be moderate quality (2 or 3🔒), and the findings suggest that FRIENDS can have a positive effect on measures of anxiety. The first three studies measured the impact of FRIENDS when used with selected participants. Barrett et al. (2001) conducted a quasi-experimental study (2🔒) that revealed significant positive effects between groups on self-ratings of anxiety and outlook. Similarly, a 2🔒 quasi-experimental study by Barrett et al. (2003) reported significant improvements on self-ratings of anxiety and outlook at post intervention and at 6-month follow-up. However, another 3🔒 RCT conducted by Hunt et al. (2009) found no significant effects at 24-months and very little significant improvements at 49-months follow-up.

A small number of 2-3🔒 studies indicate that FRIENDS has positive impact when delivered as part of a universal provision. Lowry-Webster et al. (2001) conducted an RCT (2🔒) and found significant positive effects on measures of anxiety for all learners in the intervention group, and a significant reduction of anxiety symptoms for at risk students. In a follow up study, intervention gains were maintained at 12-months (Lowry-Webster, 2003; 2🔒). In an RCT judged to be 2🔒 by Barrett and Turner (2001), the impact of teacher-led FRIENDS, psychologist-led FRIENDS and control were compared, and the results found that participants in both the intervention groups reported significant decreases in anxiety compared to the control. Similarly, in an RCT (2🔒) by Lock and Barrett (2003), reductions in self-report anxiety were significantly greater for participants in the intervention group at post-intervention and at 12-month follow-up. In follow-up to Lock and Barrett’s (2003) RCT, Barrett et al. (2006) found mixed results (2🔒). For example, reductions in anxiety were reported at 12-, 24- and 36-month follow-up for grade 6 but not grade 9. In another 2🔒 study by Barrett et al. (2005), no significant effects between groups on measures of anxiety were found post intervention, but effects were found at 12-month follow-up. In line with the majority of findings, a 3🔒 study by Essau et al. (2015) found significant positive effects on measures of anxiety through a large-scale RCT post intervention, and at 12-month follow-up, although following a Bonferroni adjustment, differences were only detected between the two groups at 12-months follow-up. The quasi-experimental study (2🔒) conducted by Matsumoto and Shimizu (2016) reported significant positive effects on measures of social phobia for girls in the intervention group compared to the control, and on measures of obsessive-compulsive disorder. However, no other significant interactions were observed.

In all of the previously mentioned studies, rated 2-3🔒, schools were the unit of randomisation rather than the students. Nevertheless, all of the studies reported baseline equivalence. However, in a number of the studies baseline equivalence was not established (Barrett et al., 2001, 2005, 2006; Lowry-Webster et al., 2001, 2003; Matsumoto & Shimizu, 2016). There were also other methodological weaknesses to consider that are likely to impact findings. For example, there was a lack of reporting on attrition (Barrett et al., 2001, 2003; Barrett & Turner, 2001; Essau et al., 2015) and, when reported, attrition levels were moderate to high (Barrett et al., 2005; Barrett & Turner, 2001; Lowry-Webster et al., 2003); student self-report measures were used exclusively; and, all except three studies were conducted by the programme developers (Essau et al., 2015; Hunt et al., 2009; Matsumoto & Shimizu, 2016). Notably, most of the studies neglect to analyse data at the appropriate level and subsequently failed to account for the effects of student clustering within the same classroom/schools, which may compromise the validity of findings (Barrett et al., 2001, 2003, 2005; Barrett & Turner 2001; Essau et al., 2015; Hunt et al., 2009; Lowry-Webster et al., 2001, 2003; Matsumoto & Shimizu, 2016). It is also important to highlight that in many of the studies that reported positive effects, the implementation of FRIENDS was by mental health professionals (Barrett et al., 2001, 2003, 2005; Essau et al. 2012; Lock & Barrett, 2003; Matsumoto & Shimizu, 2016). A 2🔒 study by Barrett and Turner (2001) indicated that both teachers and mental health professionals can deliver the FRIENDS program effectively. However, when teachers delivered the programme, results were mixed. For example, primary positive effects on anxiety were reported in three 2🔒 studies in which teachers delivered the programme (Barrett et al., 2006; Lowry-Webster et al., 2001, 2003) and no impact was found in the remaining 3🔒 study (Hunt et al., 2009).

There were a further four large-scale studies included in the review that employed multi-level modelling to analyse the data, and these were rated high quality (4🔒). Two RCTs were conducted by Miller et al. (2011). The first RCT targeted pupils with anxiety symptoms and the second took a whole-class approach. The results from both studies found no statistical differences between the intervention and control group on self, teacher, and parent-report measures of anxiety. Although the schools were the unit of randomisation and some group imbalance was reported, equivalence was established at baseline. Conversely, Stallard et al. (2014) found significant positive effects on measures of anxiety at 12-months for all students in a health-led Friends (mental health professional delivery) condition compared to school-led Friends (teacher delivery) and in a health-led Friends compared to usual provision. No differences were found between groups on self-reports of depression, nor on the teacher and parent-report measures. To mitigate the effects of randomising at school level, Stallard et al. (2014) balanced the trial groups based on key characteristics. Similarly, in a cluster RCT by Ahlen et al. (2018) no short- or long-term effects of FRIENDS on child, parent and teacher measures of anxiety and depression were found. Again, schools were the unit of randomisation and, although groups were unequal at baseline on two variables, efforts were made to account for the differences. Only minimal attrition was reported post intervention and moderate attrition at follow-up, and this was balanced across groups. Similarly, a recent large-scale cluster RCT commissioned by the EEF (Wigelsworth et al., 2018) randomised classes to conditions, and found no evidence that FRIENDS had a positive overall impact on academic attainment, and no evidence that FRIENDS has a positive impact on student ratings of anxiety and depression and teacher ratings of emotional and behavioural difficulties. Furthermore, group equivalence was established, and minimal missing data was reported (Wigelsworth et al., 2018).

Across four of the 4🔒 rated studies, results found that teacher-administered FRIENDS did not have a positive overall impact on levels of anxiety and depression in learners (Ahlen et al., 2018; Miller et al. 2011; Stallard et al., 2014). Results from a study by Stallard et al. (2014) reported similar findings and found positive overall outcomes when FRIENDS was delivered by mental health professionals.

The evidence gathered during this study suggests that Friends for Life is a promising programme that leads to reductions in anxiety. However, some of the research supporting the programme lacks methodological rigor (Higgins & O'Sullivan, 2015; Maggin & Johnson, 2014). Notably, more recent, robustly designed studies have found no-significant effects from the universal and selective implementation of FRIENDS on measures of anxiety, except when delivered by mental health professionals (see Table 13).

Table 13. *A summary of the evidence for Friends for Life (FRIENDS)*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Barrett et al. (2001)	Quasi-experimental	7–19 years, *M* age 12.42, Selective	*nintervention* = 121 *ncontrol* = 83	Effective (n r)	2🔒
Barrett & Turner (2001)	RCT	10.75 years, Grade 6, universal Australia	*nintervention* = 188 & 263 *ncontrol* = 137	Effective (n r)	2🔒
Barrett et al. (2003)	Quasi-experimental	6-19 years, *M* age 12.75, Selective, Australia	*nintervention* = 166 *ncontrol* = 154	Effective (n r)	2🔒
Barrett et al. (2005)	RCT	Grade 6 & 9, universal, Australia	*nintervention* = 423 *ncontrol* = 269	Non-significant, effective at 12-months (n r)	2🔒
Essau et al. (2012)	RCT	9-12 years, *M* age 10.91, universal, Germany	*n*intervention = 302 *n*control = 336	Effective (n r)	3🔒
Hunt et al. (2009)	RCT	Grade 7, *M* age 12.05, selective, Australia	*n* = 228 at 24 months *n = 189 at 48 months*	Non-significant, very small effect	3🔒
Lock & Barrett (2003)	RCT	Grades 6 & 9, universal, Australia	*n*intervention = 442 *n*control = 295	Effective (n r)	2🔒

Barrett et al. (2006)	RCT follow-up	Grades 7 & 10, universal, Australia	*n* = 669 at 12 months *n* = 430 at 24 months *n* = 226 at 36 months	Mixed (0.05, 0.04)	2🔒
Lowry-Webster et al. (2001)	RCT	Grades 5-7, universal, Australia	*n*intervention = *432* *n*control = 162	Effective (n r)	2🔒
Lowry-Webster et al. (2003)	RCT follow-up	Grades 5-7, universal, Australia	*n*intervention = *432* *n*control = 162	Effective (n r)	2🔒
Matsumoto & Shimizu, (2016)	Quasi-experimental	Grade 6, 11-12 years, Universal, South Africa	*n*intervention = 92 *n*control = 58	Mixed (0.03, 0.05)	2🔒
Ahlen et al. (2018)	Cluster RCT	Grades 3-4, *M* age 9.06, universal, Sweden	*n*intervention = 353 *n*control = 342	Non-significant	4🔒
Miller et al. (2011)	RCT	Grades 4-6, *M* age = 10.1, selective, Canada	*n*intervention = 65 *n*control = 126	Non-significant	4🔒
Miller et al. (2011)	RCT	Grades 4-6, *M* age 9.8, universal, Canada	*n*intervention = 141 *n*control = 112	Non-significant	4🔒
Stallard et al. (2014)	Cluster RCT	Year 4 & 5, 9-10 years, Universal, UK	*nhealth-led* = 486 *nteacher-led* = 462 *n*control = 391	Mixed	4🔒
Wigelsworth et al. (2018)	Cluster RCT	Year 5, universal, UK	*n*intervention = 1476 *n*control = 1534	Non-significant	4🔒

Only primary outcomes reported in the table, *M* age = mean age, RCT = Randomised Controlled Trial, nr = effect size not reported.

Picture Exchange Communication System (PECS)

The Picture Exchange Communication System (PECS) is a communication programme that is widely used in schools, particularly with children with autism. The current review identified seven studies related to PECS. Four of the studies used single-case multiple-baseline designs (Boesch et al., 2013; Charlop-Christy et al., 2002; Kravits et al., 2002; Tincani et al., 2006), and one was a single-case design (Ganz et al., 2004). All studies report positive effects on spontaneous communication, post intervention.

The remaining two studies were rated as low quality (1🔒) due to small sample sizes and lower-quality research designs. One was a pre-post group design, and the other a very small quasi-experimental study (Carr & Felce, 2007). Both studies report positive effects on spontaneous communication, post intervention.

There was one small RCT rated as moderate quality (2🔒), which found a significant positive effect on rates of communicative initiations with the use of PECS. However, no maintenance effects were reported, nor any significant increase in frequency of speech (Howlin et al., 2007). Despite small group sizes, this study was rated as moderate quality due to the authors' use of multi-level modelling to evaluate findings, and efforts were made to account for group differences (see Table 14).

Picture Exchange Communication System has promising evidence. We would recommend that research focuses on more robust randomised research to strengthen the promising evidence (Flippin et al., 2010; Preston & Carter, 2009).

Table 14. *A summary of the evidence for the Picture Exchange Communication System (PECS)*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Howlin et al. (2007)	RCT	*M* age 6.8 years, autistic	*nintervention = 26* *nintervention = 30* *ncontrol = 28*	Effective (n r)	2🔒

Intervention programmes and approaches with preliminary evidence of causal effect or evidence of ineffectiveness.

The following section will summarise the evidence for the intervention programmes and instructional approaches that have preliminary evidence of causal impact on a pupil outcome that is rated low quality (studies rated 0-1🔒) and those found to have no positive causal impact. See Tables 7 - 9 for the interventions that had no peer-reviewed published evidence at the time of this review.

Catch-Up Literacy

Catch-Up Literacy is a structured one-to-one intervention that uses a book-based approach to support struggling readers. This review retrieved two studies both commissioned by the EEF and judged to be moderate quality (see Table 15). Rutt (2015) identified small gains in reading following the use of Catch-Up Literacy, but the findings were not statistically significant. The remaining study found no evidence of impact on pupils' reading comprehension scores (Roy et al., 2019). Currently, there is little evidence that Catch-Up Literacy impacts positively on pupils' reading outcomes, and there is a need for more robust randomised research.

Table 15. *A summary of the evidence for Catch-Up Literacy*

Reference	Design	Population	Sample	Outcome (ES)	Quality
Rutt (2015)	RCT	Year 6 & 7	*n*intervention = 286 *n*control = 271	Non-significant (0.12)	3🔒
Roy et al. (2019)	RCT	Year 4 & 5	*n*intervention = *n* 514 control =511	No effect	3🔒

Headsprout

Headsprout is a computer assisted instructional reading programme that uses adaptive, online instruction to teach students reading and comprehension skills. This review found 14 studies that met the inclusion criteria for our review. All but one of the studies reported positive effects on reading skills from using Headsprout. However, much of the research is considered small in scale and as a result did not exceed a quality rating of 1🔒.

Three studies used a single-case multiple baseline design (Clarfield & Stoner, 2005; Cullen et al., 2014; Whitcomb et al., 2011), one was a multiple case series (Grindle et al., 2020), one was a single-case multiple probe design (Hammond & Shannon, 2015), and three were single-case pre-test post-test designs (Herring et al., 2019; Hill et al., 2015; Tyler et al., 2015). Three of these studies were conducted with learners with autism (Hill et al., 2015; Grindle et al., 2020; Whitcomb et al., 2011); three with learners that had intellectual disabilities (Cullen et al., 2014; Herring et al., 2019; Tyler et al., 2015); one with mainstream primary-aged learners at risk of reading difficulties (Hammond & Shannon, 2012); and one study was with learners diagnosed with attention deficit hyperactivity disorder (Clarfield & Stoner, 2005).

There was one small study by Kreskey and Truscott (2016) that was ranked 0🔒 for methodological quality, based on a causal comparative design (ex post facto) and non-equivalent groups. Contrary to other research in this review, this study reported negative impact from using Headsprout. The remaining studies, all rated 1🔒, found positive effects from Headsprout on reading skills. There was one small quasi-experimental study by Watkins et al. (2016) that reported positive effects on the reading skills of mainstream primary-aged pupils, although groups were non-equivalent and imbalanced. Tyler et al. (2015) conducted a small RCT with mainstream primary-aged pupils and found small, medium and large positive effects on pupils’ reading skills. However, this study reported high levels of attrition that resulted in non-equivalent groups, which weakens causal claim. A small RCT by Storey et al. (2020) also reported large positive effects on reading skills for mainstream primary-aged pupils at risk of reading failure. The remaining two small RCTs reported positive effects on reading with learners with intellectual disabilities (Ramdoss et al., 2020; Tyler et al., 2020).

Collectively, the current research suggests that Headsprout appears is an effective programme for improving the reading skills of mainstream primary-aged pupils at risk of reading failure, and for learners with intellectual disabilities and/or autism. We recommend there is a need for more robust and replicated research, across different populations to strengthen the evidence-base.

The Reading Educational Assistance Dogs (R.E.A.D) programme

R.E.A.D is a charity-based scheme that provides registered therapy dogs to schools, to support children who are reluctant readers. This review found only two group design studies based specifically on the R.E.A.D programme. The two studies that were included in this review found some evidence of impact on pupils' reading skills (Le Roux et al., 2014, 2015). However, these studies were rated low quality (1🔒) as a result of small samples and poor design. Notably, a previous review by Hall et al. (2016) found 27 studies of 'children reading to dogs' (non-specified programmes) and reported that reading to dogs may have beneficial effects on the reading environment and, subsequently, improved reading outcomes. However, much of the evidence in that review was rated low quality and R.E.A.D. was considered to have only preliminary evidence of short-term impact (Hall et al, 2016). We recommend a need for more robust and replicated research to develop the evidence-base.

Toe by Toe

Toe by Toe is a structured phonics-based reading manual for anyone who finds reading difficult. There is very little evidence for this intervention. This review found only two group design studies and both were rated low quality (0 -1🔒). Although both studies reported some positive effects on pupils' reading skills, they were very small scale and with weak designs. Therefore, there is currently no available evidence to suggest that Toe by Toe is an effective intervention, and we would recommend more robust research is needed to develop the evidence base.

Incredible Years, Dinosaur School

The Incredible Years (IY), Dina Dinosaur's Social Skills and Problem-Solving Curriculum (IY Dinosaur School) is a classroom prevention curriculum that uses dinosaur-themed materials and puppets to strengthen students' social and emotional skills. Although the IY series of intervention programmes have been the focus of a significant and sustained programme research, there is little research on the Dinosaur School Curriculum (Pidano & Allan, 2015). This review included two studies that were low quality (0 -1🔒). The first small pre-post pilot study by Hutchings et al. (2004) reported significant positive effects on measures of mental health, and non-significant gains on ratings of self-control. In a more recent small quasi-experimental study by Hutchings et al. (2012), no positive effects were found.

Currently, there is very little causal evidence available for the IY Dinosaur School programme, and we recommend more robust research is required to develop the evidence base.

Mindfulness in Schools Project, Paws b

Paws b is a school-based mindfulness curriculum for children aged 7 -11 years. This review included two studies that were both rated low quality (0 -1🔒). The first very small-scale RCT by Thomas and Atkinson (2016) found significant effects from the intervention on intentional functioning at post intervention and follow up. Another small-scale quasi-experimental study by Vickery and Dorjee (2016) revealed significant, large effects from the intervention on self-report measures of emotional affect at follow-up, but no effect on parent-report measures. Currently, there is a limited evidence base for Paws b, and we recommend more robust research is required to develop this evidence base.

Season for Growth

Seasons for Growth is a manualised grief education programme for children and young people (aged 6 to 18 years) who have experienced significant loss due to death or family breakdown. The programme promotes the normalisation of grief within the school setting. The current review identified only one single-case multiple baseline design study by Riley (2012) with 12 participants that reported positive effects on pupils’ emotional health and coping behaviours. Consequently, it is not possible to draw conclusions about the effectiveness of this intervention and more robust research is needed.

Time-out

Time-out is a behaviour change technique that can be used in multiple settings to decrease the frequency of a target behaviour using a function-based approach. This review found one study. Donaldson and Vollmer (2011) conducted a single-case reversal design and found that time-out procedures were effective at reducing problem behaviour. A review conducted by Vegas et al. (2007) did find several single-case design studies that showed Time-out as effective when dealing with attention-maintained behaviour. However, each of these studies were conducted over 20 years ago and thus outside of the date range specified for this review.

Time-out has promising findings from single-case design studies, and we would recommend more robust research designs and replication with larger number of participants to strengthen its evidence base.

Student Assistance Program (SAP)

The Student Assistance Programme (SAP) is a comprehensive model of primary prevention and early intervention that utitilises learning activities and support groups for vulnerable students across all key stages. This review found only one causal comparative design (ex post facto) study by Bindle et al. (2014) which was rated low quality (1🔒). Even though these findings from Bindle et al. (2014) were somewhat positive, there is currently insufficient evidence of the causal impact of SAP on learner outcomes; and we recommend more robust research is required to develop this evidence base.

Language for Thinking

A classroom resource that provides a clear structure to assist school staff in developing children's language from the concrete to the abstract. This review found only one low-quality pre-post small-scale study (1🔒), which reported positive effects on pupil outcomes. Currently, conclusions about the effectiveness of this resource cannot be drawn from the evidence; we recommend more robust research is required to develop this evidence base.

Social Stories

Social Stories are individualised short stories that help children and adolescents with autism spectrum disorders to better understand social situations. The current study identified 13 studies. All except one study by Malmberg et al. (2015) reported positive impact on problematic and social behaviours. All studies employed a single-case multiple baseline design (Beh-Pajooh et al., 2011; Chan & O'Reilly, 2008; Chan et al., 2011; da Silva et al., 2020; Delano & Snell, 2006; Hanrahan et al., 2020; Kim et al., 2014; Malmberg et al., 2015; O’Connor & Hayes, 2019; Sansosti & Powell-Smith, 2006; Schneider & Goldstein, 2009; Srija et al., 2019; Thompson & Johnston, 2013).

Each of the studies adds to the causal evidence; however, we recommend more robust randomised research is required to develop this evidence base.

A Narrative Intervention by Becky Shanks

Narrative interventions help to develop learners’ speaking and listening skills through the telling or retelling of stories that have specific language-related features. There is a substantial evidence-base for narrative interventions; however, in most of the literature it is not clear which intervention is being used. A search of the literature found two reviews of non-programme-specific narrative interventions. Petersen (2011) reviewed nine studies and reported moderate to large positive effects, and Favot et al. (2020) reported moderate positive effects across 24 studies. In both reviews the evidence-base was considered to be low quality due to small sample sizes and lack of robust experimental designs.

This review found only one study evaluating Becky Shanks's narrative intervention, conducted by the author (Davies et al., 2004). Although the study reported positive effects from the intervention, the study was rated low quality as a result of the small sample, the use of a pre-post design and a conflict of interest (1🔒).

Currently, conclusions about the effectiveness of this intervention cannot be drawn from the evidence and we recommend more robust research is required to develop this evidence base.

Lego® -Based Therapy

LEGO-Based Therapy is a social development programme that uses LEGO activities to support the development of a wide range of social skills within a group setting. Three studies were selected for this review and all reported positive effects on the social interactions of ASD students that used the intervention. There were two single-case design studies, of which one was a single-case reversal design (Hu et al., 2008) and the other a single-case multiple-baseline study (Levy & Dusmuir, 2020). The final study was a small-scale, pre-post (Andras, 2012), which was rated low quality (0-1🔒) due to its small scale and lack of a control comparison group. Due to this limited evidence base, it is currently not possible to draw any conclusions about the effectiveness of Lego-Based Therapy, and we recommend more robust research is required to develop this evidence base.

Results from the follow-up investigation

This section will report the results from the follow-up study that was conducted 1-year later. The follow-up aimed to investigate whether increased knowledge about the evidence for the interventions used in the cluster, presented through a summary report to school leaders, had an effect on existing provision. The data from the completed checklists and interviews are presented.

Follow-up data was received from four schools. Three of the schools (two primary and one special needs school) made no changes to their provision and continued to use the same interventions that were identified during the initial screening process. The remaining primary school (referred to as X) made changes to their provision and discontinued use of three interventions. Two of the discontinued interventions had no evidence, and one had evidence of ineffectiveness. School X continued to use other interventions with no causal evidence and ceased use of an intervention with promising evidence. Since the primary data collection, the same school had also adopted a new intervention with no causal evidence. See Table 16 for a list of the interventions in use in school X.

Table 16. *Interventions in use in primary school X at follow-up*

Interventions previously identified in the school	Causal evidence	In use	Not in use
SAFMEDS	Promising	x	
Headsprout	Preliminary	x	
Catch Up Literacy	Ineffective		x
Read Write Inc. Phonics	No evidence	x	
Seasons for Growth	Preliminary	x	
Friends for Life	Promising		x
Be Smart, Be Cool, Be Positive	No evidence	x	
Emotional Literacy Support Assistant (ELSA)	No evidence	x	
Reading Rocks	No evidence		x
Peer support	No evidence		x
Narrative Therapy Beck Shanks	Preliminary	x	
Elklan, Talking Matters	No evidence	x	
Sumdog	No evidence	x	
My Maths	No evidence	x	
Times Table Rockstars	No evidence	x	
*Talkabout	No evidence	x	

* A recently adopted intervention

Three head teachers participated in the interviews. They were asked whether access to additional information about the evidence base of interventions currently used in their school, had an effect on the way they select interventions and programmes. Two main themes were identified:

Compatibility and impact

Most of the senior leaders spoke about cost and effectiveness/impact as important factors when choosing or adjusting provision. One senior leader talked about using the information about causal-impact for the interventions, and cost, to make changes to the school provision. Other compatibility factors mentioned were the effective use of teachers' time and the ability to use interventions and teaching approaches at home alongside their compatibility with existing school systems and approaches. However, all the senior leaders cited the cost of some programmes as being a prohibitive factor. The senior leader who made some changes to provision following the presentation of the summary reports of evidence, indicated that the school had continued to use interventions because of their compatibility with the school and its teaching practice. For example, interventions could be used at home.

> "Yeah, hugely. So as a school, we've started looking at research for various areas, and what the research says. For the interventions, we've now got an impact intervention costs for everything to do with literacy, and maths, that will be done in wellbeing, too. So the interventions that we currently use have been costed down to the pupil, cost per pupil, and time versus cost, and the effect and evidence, size of those interventions. Anything that was either to higher cost or ineffective has been pulled. And the money has been transferred, and the time of the teaching assistance has been transferred into one that does work or have an impact."

> "Very few interventions have disappeared completely, partly due to the fact that we've streamlined it a while ago. Some are being used, partly, but most of those interventions that were on the list have continued and will continue to go forward. We don't have homework per se, My Maths is done as a weekly homework based on the topic that they've been doing in school that week, or that two weeks. Sum Dog and TT Rockstars are there for them to use at home should they wish."

> "What it's made me really look at is cost and impact? Because, if I take Catch-Up reading the cost of a teaching assistant to run that was about £20,000, and the impact was no better than the Headsprout programme, which basically costs hardly anything, and it can be disseminated wider across the school."

> "And it's also made me really think about… there's lots of companies always trying to push things, always trying to push this intervention and this and it's just about that whole taking a step back, looking at the impact and looking at the cost."

Greater awareness of the evidence

All three senior leaders referred to a greater awareness of the need to assess the impact of interventions. Two referred to having developed a greater awareness of the evidence for interventions and programmes following their engagement with the current study and spoke about using the evidence to improve the quality of decision making.

> "It's just made me really think about interventions and look more at what other programmes are out there that have a broader evidence base. I think it's just made me more aware of looking at the evidence."
>
> "We have very much more awareness of interventions having an evidence base. So we would look to that first. We really do think now before we even send somebody on training for something. Has this got evidence behind it? Is it proven to be worthwhile to do in school? Previously we would send somebody on a course, after it had come out on an email. We'd send somebody along, we'd do the intervention because somebody said other teachers came back and liked it, and that was it, really. I think we do think an awful lot more about the worthwhileness of things now, definitely. Yeah."

Discussion

Very little is known about the range of intervention programmes and approaches used in schools, or the evidence supporting them. Several published studies report on teachers and senior leaders' engagement with and use of research (e.g., Coldwell et al., 2017; Nelson et al., 2017; Walker et al., 2019). To our knowledge, this study was the first systematic attempt to assess the range of interventions used in a large school cluster in the UK, and to establish which of these programmes and instructional approaches have evidence of causal effect on pupil outcomes.

Our results found that 138 interventions were in use across the 10 schools and this number was judged to be high. However, the school cluster in this study is in an area of high deprivation and reported a higher number of ALN pupils than typically reported elsewhere in the UK (DfE, 2019; Welsh Government, 2019). Our results suggest that schools are using a large number of interventions. The use of a large number of interventions has obvious implications on school budgets, staff time. Additionally, the use of many interventions in one setting impacts significantly on the likelihood any one intervention will be implemented with fidelity.

In this review we categorised the interventions into three main SEN areas of need: cognition and learning; social, emotional and mental health; and, communication and interaction, which reflects the most common types of support reported by Welsh Government (2019). There were 55 interventions in the cognition and learning category, 64 interventions in the social, emotional and mental health category (and a little under half of these were programmes to teach social skills and emotional resilience) and 19 interventions in the communication and interaction category.

Although this distribution of interventions according to SEN category in part reflects the focus of school funding on provision for teaching and learning approaches and learners' well-being, the relatively small number of interventions focused on improving learners' communication is notable, but might just reflect the paucity of available provision in this area. The notable feature of this research is the large number of interventions in use across the cluster (138) consisting of 70 unique programmes and/or approaches. The two schools with the highest number of interventions (24 and 25) each have speech and language provision which might, in part, explain the need for additional interventions. The school with the fewest interventions (4) is a Welsh medium school, which may reflect the paucity of resources and interventions available in the Wels language and an urgent need to develop Welsh medium interventions across all intervention areas. The two secondary schools have fewer interventions than many of the primary schools, which may suggest, as did reports from the DfE (2017) that there is a more limited evidence-base for educators in secondary schools. An important feature of the findings is the lack of relationship between the number of ALN pupils in each school and the number of interventions employed, and the high number of social, emotional and mental health interventions used in some schools. It is likely that the high number of interventions recorded in some schools is the indirect result of cumulative additional funding designed to help disadvantaged pupils ‘catch up’.

Within the context of our focused review, we also found that many of these interventions lacked robust empirical support. The results from our review showed that only 30 per cent (42) of interventions used across the cluster had positive evidence of impact on pupil outcomes. However, of these, only 11 per cent (15) had evidence of effectiveness that was rated moderate to high quality and considered promising, and 19 per cent (27) had preliminary evidence, rated low quality and in need of more research evidence. Notably, only FRIENDS and Accelerated Reader had high quality replicated evidence. The results also showed that 67 per cent (92) of all interventions used across the cluster had no published evidence at the time of this review. The remaining three per cent of interventions were shown to be ineffective. Furthermore, a large number of the interventions, in the social, emotional and mental health category were non-specific and as a result had no evidence. For example, most of the schools implemented nurture groups and did not implement a specific programme or follow a defined nurture approach or protocol. Obviously, we could not make any assessment of the effectiveness of these approaches.

This is the first study to empirically report the evidence base for the interventions used in a cluster of schools, and the results suggest that a very small number of these interventions have causal evidence for positive impact on pupil outcomes. The findings from our study supports findings by Gorard et al. (2020) and suggests that even when presented with summary reports of the evidence, not all schools use this information to make more informed decisions about school provision. Instead of accessing external evidence, it is common for teachers to turn to anecdotal evidence, including advice from colleagues and to programmes that are compatible with current practice as opposed to those with the stronger evidence of positive impact on learner outcomes (Brown & Greany, 2018; Nelson et al., 2017; Walker et al., 2019). Our data also suggest that even if school leaders are interested in reviewing the evidence of interventions, generally there is a paucity of robust education intervention research and much more resources and energy will be required to build robust evidence across almost every area of education. Previous research has reported that barriers to evidence-into-use include lack of time, underdeveloped research skills, negative attitudes and opinions of evidence, and, a lack of relevant, accessible and usable evidence may be preventing teachers from using the external evidence (Cain, 2016; Gorard, 2020; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018; Walker et al., 2019). The findings from this study support these previous findings, and further suggest that the lack of an open-source evidence base for academic and well-being interventions is likely to be a contributing factor leading to the accumulation of numerous interventions and strategies, many of which lack a strong evidence base.

Limitations

This review has some limitations that should be considered when interpreting the results. First, although the study selection criteria and search strategy allowed for the capture of causal evidence it cannot be assumed that all the available evidence was included. As a result of resource constraints and inclusion criteria our review is not exhaustive, and we recognise that some research might not be included. We would recommend that future research could extend the review process to be more expansive. This review also chose to exclude the grey literature and, whilst the first author was aware of the negative effects of publication bias, it was felt from a preliminary scan that the grey literature would not improve the quality of the results. Also, resource constraints prevented a second reviewer carrying out an independent screening of a sample of papers.

It is also important to highlight that this paper is not advocating for any of the programmes or approaches included in this review or rejecting the interventions with little to no evidence. Our aim is to objectively present the available evidence to the participating schools, and inform the wider education system of the need to consider the quality of the research evidence when considering school interventions.

Implications for the school cluster

We hope the findings from this study will inform provision and enable the schools to move towards a more evidence-informed approach to provision, which is informed by the availability of higher quality research evidence. Doing so, will help to ensure scarce resources are spent more efficiently on fewer interventions with a stronger evidence base for impact on pupil outcomes. When undertaking reviews such as this, schools might also consider using approaches from health economics such as Programme Budgeting Marginal Analysis (PMBA) to further evaluate and refine the range of interventions in terms of need, resource implication, cost effectiveness, and social validity. For schools that choose to continue to use programmes and approaches that lack evidence, we would encourage them to start the process of evaluation of impact. This could be achieved initially through the use of school-level data, action research projects or through collaborative research projects with higher education institutions.

Implications for policy

The outcomes of this study support previous findings by Schaik et al. (2018) and indicate that government agencies need to work more closely with schools and researchers to ensure a greater number of robust and relevant studies are conducted to strengthen the existing evidence base for interventions. Policymakers and researchers should work together with funders to create an open access database of evaluation findings for commonly used interventions and programmes.

Directions for research

Evidence from this review provides some insight into the programmes that would most benefit investigation. Conclusions drawn from a review by Schaik et al. (2018) and Gorard et al (2020) report that more collaborations and partnerships between researchers and teachers are necessary to facilitate more effective evidence-use in schools. We suggest that the findings from this study could be used to inform future collaborative projects to ensure researchers work more closely with schools to develop a more robust and relevant evidence base for interventions. What we have attempted to do here within the constraints of this project, is to develop a rapid review process for examining the evidence base for interventions. Researchers should now consider replicating the current study in other regions across the UK to begin to create a database of the commonly used programmes and approaches. In addition, and in agreement with Gorard et al. (2020), our results indicate that more empirical research needs to be conducted to understand the best way to facilitate use of the external evidence in schools, by decision makers, to inform provision.

Conclusions

The present study provides important new information on the range and evidence-base of interventions in use in a cluster of schools. The findings from this review suggest that many schools adopt a large number of interventions that presently have a limited evidence-base, with few having evidence of positive causal impact on learner outcomes. The results from this study can be used to inform provision at school level and help the schools and cluster move towards a more evidence-informed approach. Governments, education funders and the research community should use these findings as a starting point for creating a robust and relevant repository to help schools identify more effective interventions and to develop the evidence base and to inform future collaborative research projects.

References from the rapid systematic review

Accelerated Reader

What Works Clearing House. (2007). *Accelerated Reader: WWC Intervention Report*. Retrieved from ERIC Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-reading-renaissance-what-works/docview/62020030/se-2?accountid=14874

What Works Clearing House. (2009). *Accelerated Reader: WWC Intervention Report*. Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-what-works-clearinghouse/docview/61823212/se-2?accountid=14874

What Works Clearing House. (2010). *Accelerated Reader: WWC Intervention Report*. Retrieved from ERIC Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-tm-what-works-clearinghouse/docview/754911655/se-2?accountid=1487

Brem, S. K., Husman, J., & Duggan, M. A. (2005). Findings from a three-year study of Reading Renaissance in a Title I urban elementary school. *Academic Exchange Quarterly*, 9(3), 355-359.

Brown, K. L. (2010). *The impact of accelerated reader (AR) program on students' MCT reading scores* Available from Social Science Premium Collection. (889926355; ED522354). Retrieved from http://ezproxy.bangor.ac.uk/login?url=https://www.proquest.com/dissertations-theses/impact-accelerated-reader-ar-program-on-students/docview/889926355/se-2?accountid=14874

Bryant, W. E. (2008). Effect of the Accelerated Reader program on academic achievement (Doctoral dissertation, Northcentral University). *Dissertation Abstracts International*, 69(02A), 90–443.

Bullock, J. C. (2005). Effects of the Accelerated Reader on the reading performance of third, fourth, and fifth-grade students in one western Oregon elementary school (Doctoral dissertation, University of Oregon). *Dissertation Abstracts International*, 66(07A), 56–2529.

Chase, E., Goodin, P., & Nichols, W. R. (2010). Evaluation of the accelerated reader program in chesapeake, VA, public schools. *ERS Spectrum*, 28(1), 25-32. Retrieved from https://www-proquest-com.ezproxy.bangor.ac.uk/scholarly-journals/evaluation-accelerated-reader-program-chesapeake/docview/968116421/se-2?accountid=14874

Friesen, C. (2004). *Improving reading in grade three students*. Wisconsin Rapids, WI: Renaissance Learning.

Griffin, T. F. (2000). A causal comparative study on the effects of Accelerated Reader (Doctoral dissertation, University of North Carolina at Charlotte). *Dissertation Abstracts International*, 61(08A), 99–3106

Hagerman, T. E. (2003). A quasi-experimental study on the effects of Accelerated Reader at middle school (Doctoral dissertation, University of Oregon). *Dissertation Abstracts International*, 64(06A), 124–2027.

Hart, S. S. (2007). Accelerated Reader in a primary school: An evaluation of time spent on classroom implementation and student achievement (Doctoral dissertation, Capella University). *Dissertation Abstracts International*, 68(04A), 122–1384.

Holloway, A. (2007). The effects of Accelerated Reader on the attitudes and reading habits of first grade students in a mid-southern state (Doctoral dissertation, Union University). *Dissertation Abstracts International*, 68(11A), 108–4610.

Holmes, C. T., & Brown, C. L. (2003). A controlled evaluation of a total school improvement process, *School Renaissance*. Athens, GA: University of Georgia, Department of

Educational Administration

Huang, S., PhD. (2012). A mixed method study of the effectiveness of the accelerated reader program on middle school students' reading achievement and motivation. *Reading Horizons, 51*(3), 229-246. Retrieved from https://www.proquest.com/scholarly-journals/mixed-method-study-effectiveness-accelerated/docview/921280084/se-2?accountid=14874

Hunter, T. A. (2013). *The effect of accelerated reader on reading scores in a rural school district.* Walden University. Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/dissertations-theses/effect-accelerated-reader-on-reading-scores-rural/docview/1697500399/se-2?accountid=14874

Johnson, R. A., & Howard, C. A. (2003). The effects of the accelerated reader program on the reading comprehension of pupils in grades three, four, and five. *The Reading Matrix: An International Online Journal, 3*(3) Retrieved from https://www.proquest.com/scholarly-journals/effects-accelerated-reader-program-on-reading/docview/85631356/se-2?accountid=14874

Kohel, P. R. (2003). Using Accelerated Reader: Its impact on the reading levels and Delaware state testing scores of 10th grade students in Delaware’s Milford High School (Doctoral dissertation, Wilmington College). *Dissertation Abstracts International*, 63(10A), 106–3507.

Kyllo, A. (2004). Does Accelerated Reader have positive and motivational effects on student reading levels and student attitude toward reading? In T. F. Sherman, & M. Lundquist (Eds.), Winona state university anthology of K-12 language arts action research.

Mallette, M. H., Henk, W. A., & Melnick, S. A. (2004). The influence of accelerated reader on the affective literacy orientations of intermediate grade students: JLR. *Journal of Literacy Research, 36*(1), 73-84. Retrieved from https://www.proquest.com/scholarly-journals/influence-accelerated-reader-on-affective/docview/231938693/se-2?accountid=14874

Melton, C. M., Smothers, B. C., Anderson, E., Fulton, R., Replogle, W. H., & Thomas, L. (2004). A study of the effects of the accelerated reader program on fifth grade students' reading achievement growth. *Reading Improvement*, *41*(1), 18-23. Retrieved from https://www.proquest.com/scholarly-journals/study-effects-accelerated-reader-program-on-fifth/docview/85594943/se-2?accountid=14874

McGlinn, J. M., & Parrish, A. (2002). Accelerating ESL students' reading progress with accelerated reader. *Reading Horizons, 42*(3), 175-189. Retrieved from https://www.proquest.com/scholarly-journals/accelerating-esl-students-reading-progress-with/docview/236424173/se-2?accountid=14874

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6(7), e1000097.

Nichols, J. S. (2013). Accelerated reader and its effect on fifth-grade students' reading comprehension Available from ERIC. (1697491793; ED553041). Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/dissertations-theses/accelerated-reader-effect-on-fifth-grade-students/docview/1697491793/se-2?accountid=14874

Nunnery, J. A., & Ross, S. M. (2007). The effects of the school renaissance program on student achievement in reading and mathematics. *Research in the Schools, 14*(1), 40-59. Retrieved from https://www.proquest.com/scholarly-journals/effects-school-renaissance-program-on-student/docview/211015676/se-2?accountid=14874

Nunnery, J. A., Ross, S. M., & Goldfeder, E. (2003). The effect of School Renaissance on TAAS scores in the McKinney ISD. Memphis, TN: *Center for Research in Educational Policy*

Nunnery, J. A., Ross, S. M., & McDonald, A. (2006). A randomized experimental evaluation of the impact of accelerated reader/reading renaissance implementation on reading achievement in grades 3 to 6. *Journal of Education for Students Placed at Risk, 11*(1), 1-18. http://dx.doi.org/10.1207/s15327671espr1101_1

Ostrom, J. (2007). A study of reading achievement of students participating in the Accelerated Reader program

Pavonetti, L. M., Brimmer, K. M., & Cipielewski, J. F. (2002). Accelerated reader: What are the lasting effects on the habits of middle school students exposed to accelerated reader in elementary grades? *Journal of Adolescent & Adult Literacy, 46*(4), 300-311. Retrieved from https://www.proquest.com/scholarly-journals/accelerated-reader-what-are-lasting-effects-on/docview/216914879/se-2?accountid=14874

Putman, S. M. (2005). Computer-based reading technology in the classroom: The affective influence of performance contingent point accumulation on 4th grade students. *Reading Research & Instruction*, 45(1), 19–38

Rodriguez, S. (2007). The accelerated reader programs' relationship to student achievement on the english-language arts california standards test. *The Reading Matrix: An International Online Journal, 7*(3), 191-205. Retrieved from https://www.proquest.com/scholarly-journals/accelerated-reader-programs-relationship-student/docview/85666061/se-2?accountid=14874

Ross, S. M., Nunnery, J., & Goldfeder, E. (2004). *A randomized experiment on the effects of Accelerated Reader/Reading Renaissance in an urban school district: Preliminary evaluation report*. Memphis, TN: The University of Memphis, Center for Research in Educational Policy

Ross, S., & Nunnery, J. (2005). *The effects of School Renaissance on student achievement in two Mississippi school districts: A longitudinal quasi-experimental study*. Memphis, TN: Center for Research in Educational Policy and Education Innovations.

Rudd, P., & Wade, P. (2006). *Evaluation of" Renaissance Learning" Mathematics and Reading Programs in UK Specialist and Feeder Schools*. Final Report, 2006. National Foundation for Educational Research.

Samuels, S. J., Lewis, M., Wu, Y., Reininger, J., & Murphy, A. (2003). *Accelerated Reader vs. non-Accelerated Reader: How students using the Accelerated Reader outperformed the control condition in a tightly controlled experimental study*. The University of Minnesota, Department of Education Psychology.

Topping, K. J., & Fisher, A. M. (2001). *Accelerated Reader: UK Pilot, 1999-2000*. Summary Report.

Topping, K. J., & Fisher, A. M. (2003). Computerised formative assessment of reading comprehension: Field trials in the U.K. *Journal of Research in Reading*, 26(3), 267–279.

Shannon, L. C., Styers, M. K., Wilkerson, S. B., & Peery, E. (2015). Computer-assisted learning in elementary reading: A randomized control trial. *Computers in the Schools*, *32*(1), 20-34.

Siddiqui, N., Gorard, S., & See, B. H. (2016). Accelerated Reader as a literacy catch-up intervention during primary to secondary school transition phase. *Educational Review*, *68*(2), 139-154. http://doi:10.1080/00131911.2015.1067883

What Works Clearing House (2007) *Accelerated Reader: WWC intervention report*, IES, Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-reading-renaissance-what-works/docview/62020030/se-2?accountid=14874

What Works Clearing House (2009*). Accelerated reader: WWC intervention report*, IES, Retrieved from ERIC Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-what-works-clearinghouse/docview/61823212/se-2?accountid=14874

What Works Clearing House (2010). *Accelerated reader: WWC intervention report*, IES, Retrieved from https://search-proquest-com.ezproxy.bangor.ac.uk/reports/accelerated-reader-tm-what-works-clearinghouse/docview/754911655/se-2?accountid=14874

Catch-Up Literacy

Roy. P, Rutt. S, Buchanan. E, Rennie. C, Martin. K and Walker. F (2019) *Catch Up Literacy Evaluation report and executive summary.* Retrieved from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Catch_Up_Literacy.pdf"

Rutt (2015) *Catch Up Literacy Evaluation Report and Executive Summary*. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/EEF_Project_Report_CatchUpLiteracy.pdf

Friends for Life

Ahlen, J., Hursti, T., Tanner, L., Tokay, Z., & Ghaderi, A. (2018). Prevention of anxiety and depression in Swedish school children: a cluster-randomized effectiveness study. *Prevention science*, *19*(2), 147-158.

Barrett, P. M., Farrell, L. J., Ollendick, T. H., & Dadds, M. (2006). Long-term outcomes of an Australian universal prevention trial of anxiety and depression symptoms in children and youth: an evaluation of the friends program. *Journal of clinical child and adolescent psychology*, *35*(3), 403-411. http://doi:10.1207/s15374424jccp3503_5

Barrett, P. M., Lock, S., & Farrell, L. J. (2005). Developmental differences in universal preventive intervention for child anxiety. *Clinical Child Psychology & Psychiatry*, 10, 539-555. http://doi:10.1177/1359104505056317

Barrett, P. M., Sonderegger, R., & Sonderegger, N. L. (2001). Evaluation of an anxiety-prevention and positive-coping program (FRIENDS) for children and adolescents of non-English-speaking background. *Behaviour Change*, 18(2), 78–91.

Barrett, P. M., Sonderegger, R., & Xenos, S. (2003). Using FRIENDS to combat anxiety and adjustment problems among young migrants to Australia: A national trial. *Clinical Child Psychology and Psychiatry*, *8*(2), 241-260.

Barrett, P. M., & Turner, C. (2001). Prevention of anxiety symptoms in primary school children: Preliminary results from a universal school-based trial. *British Journal of Clinical Psychology*, 40, 399–410.

Bernstein, G. A., Layne, A. E., Egan, E. A., & Tennison, D. M. (2005). School-based interventions for anxious children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*(11), 1118-1127.

Briesch, A. M., Hagermoser Sanetti, L. M., & Briesch, J. M. (2010). Reducing the Prevalence of Anxiety in Children and Adolescents: An Evaluation of the Evidence Base for the FRIENDS for Life Program. *School Mental Health*, 2(4), 155–165. http://doi.org/10.1007/s12310-010-9042-5

Cooley, M. R., Boyd, R. C., & Grados, J. J. (2004). Feasibility of an anxiety preventive intervention for community violence exposed African-American children. *Journal of Primary Prevention*, *25*(1), 105-123.

Essau, C. A., Conradt, J., Sasagawa, S., & Ollendick, T. H. (2012). Prevention of anxiety symptoms in children: Results from a universal school-based trial. *Behavior therapy*, *43*(2), 450-464. http://doi:10.1016/j.beth.2011.08.003

Fjermestad, K. W., Wergeland, G. J., Rogde, A., Bjaastad, J. F., Heiervang, E., & Haugland, B. S. M. (2020). School-based targeted prevention compared to specialist mental health treatment for youth anxiety. *Child and Adolescent Mental Health*, *25*(2), 102-109. doi:10.1111/camh.12366

Gallegos, J., Linan-Thompson, S., Stark, K., & Ruvalcaba, N. (2013). Preventing childhood anxiety and depression: Testing the effectiveness of a school based program in Mexico. *Psicología Educativa*, *19*(1), 37-44.

Higgins, E., & O’Sullivan, S. (2015). “What Works”: systematic review of the “FRIENDS for Life” programme as a universal school-based intervention programme for the prevention of child and youth anxiety. *Educational Psychology in Practice*, 31(4), 424-438.

Hunt, C., Andrews, G., Sakashita, C., Crino, R., & Erskine, A. (2009). Randomized controlled trial of an early intervention programme for adolescent anxiety disorders. *Australian & New Zealand Journal of Psychiatry*, *43*(4), 300-304.

Liddle, I., & Macmillan, S. (2010). Evaluating the FRIENDS programme in a Scottish setting. *Educational Psychology in Practice*, *26*(1), 53-67.

Lock, S., & Barrett, P. M. (2003). A longitudinal study of developmental differences in universal preventive intervention for child anxiety. *Behaviour Change*, 20, 183–199.

Lowry-Webster, H. M., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour change*, *18*(1), 36-50.

Lowry-Webster, H. M., Barrett, P. M., & Lock, S. (2003). A universal prevention trial of anxiety symptomology during childhood: Results at 1-year follow-up. *Behaviour change*, *20*(1), 25-43.

Maggin, D. M., & Johnson, A. H. (2014). A meta-analytic evaluation of the FRIENDS program for preventing anxiety in student populations. *Education and Treatment of*

Children, 277-306.

Matsumoto, Y., & Shimizu, E. (2016). The FRIENDS Cognitive Behavioral Program in Japanese schools: An examination of the treatment effects. *School Psychology International, 37*(4), 397-409. doi;10.1177/0143034316649639

Miller, L. D., Laye-Gindhu, A., Liu, Y., March, J. S., Thordarson, D. S., & Garland, E. J. (2011). Evaluation of a preventative intervention for child anxiety in two randomized attention-control trials. *Behaviour Research and Therapy,* 49, 1-9. doi:10.1016/j. brat.2011.02.00.

Mostert, J., & Loxton, H. (2008). Exploring the effectiveness of the FRIENDS program in reducing anxiety symptoms among South African children. *Behaviour Change, 25*(2), 85-96. doi:10.1375/bech.25.2.85

St. Onge, J. R., Stephenson, R., & Senthil Kumar, B. (2017). Validation of the FRIENDS anxiety prevention program for children in Canada. *Canadian Journal of Community Mental Health, 35*(3), 25-40.

Rodgers, A., & Dunsmuir, S. (2015). A controlled evaluation of the 'FRIENDS for Life' emotional resiliency programme on overall anxiety levels, anxiety subtype levels and school adjustment. *Child and Adolescent Mental Health, 20*(1), 13-19.

Rose, H., Miller, L., & Martinez, Y. (2009). "FRIENDS for life": The results of a resilience-building, anxiety-prevention program in a Canadian elementary school. *Professional School Counseling, 12*(6), 2156759X0901200612.

Schoenfeld, N. A., & Mathur, S. R. (2009). Effects of cognitive-behavioral intervention on the school performance of students with emotional or behavioral disorders and anxiety. *Behavioral Disorders, 34*(4), 184-195.

Siu, A. F. (2007). Using FRIENDS to combat internalizing problems among primary school children in Hong Kong. *Journal of Evidence-Based Psychotherapies, 7*(1), 11.

Stallard, P., Simpson, N., Anderson, S., Carter, T., Osborn, C., & Bush, S. (2005). An evaluation of the FRIENDS programme: A cognitive behavior therapy intervention to promote emotional resilience. *Archives of Disease in Childhood*, 90, 1016–1019

Stallard, P., Simpson, N., Anderson, S., & Goddard, M. (2008). The FRIENDS emotional health prevention programme: 12 month follow-up of a universal UK school based trial. *European Journal of Child and Adolescent Psychiatry*, 17, 283–289.

Stallard, P., Skryabina, E., Taylor, G., Phillips, R., Daniels, H., Anderson, R. & Simpson, N. (2014). Classroom-based cognitive behaviour therapy (FRIENDS): a cluster randomised controlled trial to Prevent Anxiety in Children through Education in Schools (PACES). *The Lancet Psychiatry*, 1(3), 185– 192. http://doi.org/10.1016/S2215-0366(14)70244-5

Wiglesworth et al (2018). *FRIENDS for life Evaluation report and executive summary*. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/FRIENDS.pdf

Headsprout

Clarfield, J., & Stoner, G. (2005). The effects of computerized reading instruction on the academic performance of students identified with ADHD. *School Psychology Review*, *34*(2), 246-254.

Cullen, J. M., Alber-Morgan, S. R., Schnell, S. T., & Wheaton, J. E. (2014). Improving reading skills of students with disabilities using Headsprout comprehension. *Remedial and Special Education*, *35*(6), 356-365.

Grindle, C., Kurzeja, O., Tyler, E., Saville, M., Hughes, J. C., Hastings, R. P., & Brown, F. J. (2020). Teaching Children with Autism Reading Comprehension Skills using Online Reading Instruction: Preliminary Evaluation of Headsprout Reading Comprehension®. *Journal of International Special Needs Education, 23*(1), 1-12.

Hammond, S. S. (2012). *The effects of the Headsprout early reading program on the literacy skills and on-task behavior of at-risk urban kindergarten students*. The Ohio State University.

Herring, E., Grindle, C., & Kovshoff, H. (2019). Teaching early reading skills to children with severe intellectual disabilities using Headsprout Early Reading. *Journal of Applied Research in Intellectual Disabilities, 32*(5), 1138-1148.

Hill, D. A., & Flores, M. M. (2015). A Preliminary Investigation of the Benefits of Computer-Aided Instruction in Reading Decoding for Students with Autism Spectrum Disorder and Other Developmental Disabilities. *Journal of the American Academy of Special Education Professionals*, *72*, 82.

Kreskey, D. D., & Truscott, S. D. (2016). Is computer-aided instruction an effective tier-one intervention for kindergarten students at risk for reading failure in an applied setting?. *Contemporary School Psychology*, *20*(2), 142-151.

Ramdoss, S. T., Harbison, A. L., & Lee, K. E. (2020). Computer-assisted instruction appears to improve reading skills in students with intellectual disabilities, despite challenges with research design to establish efficacy. *Evidence-Based Communication Assessment and Intervention*, *14*(4), 189-192.

Rigney, A. M., Hixson, M. D., & Drevon, D. D. (2020). Headsprout: A Systematic Review of the Evidence. *Journal of Behavioral Education*, *29*(1), 153-167.

Roberts-Tyler, E. J., Hughes, J. C., & Hastings, R. P. (2020). Evaluating a computer-based reading programme with children with Intellectual Disabilities: feasibility and pilot research. *Journal of Research in Special Educational Needs*, 20(1), 14-26.

Storey, C., McDowell, C., & Leslie, J. C. (2017). Evaluating the efficacy of the Headsprout© reading program with children who have spent time in care. *Behavioral Interventions*, 32(3), 285-293.

Tyler, E. J., Hughes, J. C., Beverley, M., & Hastings, R. P. (2015). Improving early reading skills for beginning readers using an online programme as supplementary instruction. *European Journal of Psychology of Education*, *30*(3), 281-294.

Tyler, E. J., Hughes, J. C., Wilson, M. M., Beverley, M., Hastings, R. P., & Williams, B. M. (2015). Teaching early reading skills to children with intellectual and developmental disabilities using computer-delivered instruction: A pilot study. *Journal of International Special Needs Education*, *18*(1), 1-11.

Watkins, R. C., Hulson-Jones, A., Tyler, E., Beverley, M., Carl Hughes, J., & Hastings, R. P. (2016). Evaluation of an online reading programme to improve pupils' reading skills in primary schools: Outcomes from two implementation studies. *Cylchgrawn Addysg Cymru/Wales Journal of Education*, 18(2), 81-104.

Whitcomb, S. A., Bass, J. D., & Luiselli, J. K. (2011). Effects of a computer-based early reading program (Headsprout®) on word list and text reading skills in a student with autism. *Journal of Developmental and Physical Disabilities*, *23*(6), 491-499.

Incredible Years, Dinosaur School

Hutchings, J., Bywater, T., Gridley, N., Whitaker, C. J., Martin-Forbes, P., & Gruffydd, S. (2012). The incredible years therapeutic social and emotional skills programme: A pilot study. *School Psychology International*, 33(3), 285-293.

Hutchings, J., Lane, E., Owen, R. E., & Gwyn, R. (2004). The introduction of the Webster-

Stratton classroom dinosaur school programme in Gwynedd, North Wales: A pilot study. *Educational and Child Psychology*, *21*(4), 4-15.

Pidano, A. E., & Allen, A. R. (2015). The Incredible Years series: A review of the independent research base. *Journal of Child and Family Studies*, 24(7), 1898-1916.

Language for Thinking

Benner, G. J., Ralston, N. C., & Feuerborn, L. (2012). The effect of the language for thinking program on the cognitive processing and social adjustment of students with emotional and behavioral disorders. *Preventing School Failure: Alternative Education for Children and Youth,* 56(1), 47-54.

Lego-Based Therapy

Andras, M. (2012). The value of LEGO® therapy in promoting social interaction in primary-aged children with autism. *Good Autism Practice*, *13*(2), 17-24.

Hu, X., Zheng, Q., & Lee, G. T. (2018). Using peer-mediated LEGO® play intervention to improve social interactions for Chinese children with autism in an inclusive setting. *Journal of Autism and Developmental Disorders*, *48*(7), 2444-2457.

Levy, J., & Dunsmuir, S. (2020). Lego therapy: Building social skills for adolescents with an autism spectrum disorder. *Educational and Child Psychology, 37*(1), 58-83.

Narzisi, A., Sesso, G., Berloffa, S., Fantozzi, P., Muccio, R., Valente, E., ... & Masi, G. (2021). Could You Give Me the Blue Brick? LEGO®-Based Therapy as a Social Development Program for Children with Autism Spectrum Disorder: A Systematic Review. *Brain Sciences*, 11(6), 702.

MISProject Paws b

Carsley, D., Khoury, B., & Heath, N. L. (2018). Effectiveness of mindfulness interventions for mental health in schools: A comprehensive meta-analysis. *Mindfulness*, 9(3), 693-707.

Thomas, G., & Atkinson, C. (2016). Measuring the effectiveness of a mindfulness-based intervention for children's attentional functioning. *Educational & Child Psychology*, *33*(1), 51-64.

Vickery, C. E., & Dorjee, D. (2016). Mindfulness training in primary schools decreases negative affect and increases meta-cognition in children. *Frontiers in Psychology*, 6, 2025.

A Narrative Intervention by Becky Shanks

Davies, P., Shanks, B., & Davies, K. (2004). Improving narrative skills in young children with delayed language development. *Educational review*, *56*(3), 271-286.

Petersen, D. B. (2011). A systematic review of narrative-based language intervention with children who have language impairment. *Communication Disorders Quarterly*, 32(4), 207-220.

Favot, K., Carter, M., & Stephenson, J. (2020). The effects of oral narrative intervention on the narratives of children with language disorder: A systematic literature review. *Journal of Developmental and Physical Disabilities*, 1-48.

Picture Exchange Communication System (PECS)

Boesch, M. C., Wendt, O., Subramanian, A., & Hsu, N. (2013). Comparative efficacy of the Picture Exchange Communication System (PECS) versus a speech-generating device: Effects on requesting skills. *Research in Autism Spectrum Disorders*, *7*(3), 480-493.

Carr, D., & Felce, J. (2007). The effects of PECS teaching to Phase III on the communicative interactions between children with autism and their teachers. *Journal of Autism and Developmental Disorders*, 37(4), 724-737.

Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the picture exchange communication system (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavior Analysis*, *35*(3), 213-231.

Flippin, M., Reszka, S., & Watson, L. R. (2010). Effectiveness of the Picture Exchange Communication System (PECS) on communication and speech for children with autism spectrum disorders: A meta-analysis. *American Journal of Speech-Language Pathology*, 19, 178-195

Ganz, J. B., & Simpson, R. L. (2004). Effects on communicative requesting and speech development of the picture exchange communication system in children with characteristics of autism. *Journal of Autism and Developmental Disorders*, *34*(4), 395-409.

Howlin, P., Gordon, R. K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of Picture Exchange Communication System (PECS) training for teachers of children with autism: a pragmatic, group randomised controlled trial. *Journal of Child Psychology and Psychiatry*, *48*(5), 473-481.

Kravits, T. R., Kamps, D. M., Kemmerer, K., & Potucek, J. (2002). Brief report: Increasing communication skills for an elementary-aged student with autism using the picture exchange communication system. *Journal of Autism and Developmental Disorders*, *32*(3), 225-230.

Magiati, I., & Howlin, P. (2003). A pilot evaluation study of the Picture Exchange Communication System (PECS) for children with autistic spectrum disorders. *Autism*, *7*(3), 297-320.

Preston, D., & Carter, M. (2009). A review of the efficacy of the picture exchange communication system intervention. *Journal of Autism and Developmental Disorders*, *39*(10), 1471-1486.

Tincani, M., Crozier, S., & Alazetta, L. (2006). The Picture Exchange Communication System: Effects on manding and speech development for school-aged children with autism. *Education and Training in Developmental Disabilities*, 177-184.

The Reading Educational Assistance Dogs (R.E.A.D) programme

Hall, S. S., Gee, N. R., & Mills, D. S. (2016). Children reading to dogs: A systematic review of the literature. *PloS One*, *11*(2), e0149759.

Le Roux, M. C., Swartz, L., & Swart, E. (2014, December). The effect of an animal-assisted reading program on the reading rate, accuracy and comprehension of grade 3 students: A randomized control study. *Child & Youth Care Forum* 43 (6), 655-673.

Le Roux, M., Swartz, L., & Swart, E. (2015). The effect of an animal-assisted reading programme on word recognition and spelling skills of grade 3 children. *Tydskrif vir Geesteswetenskappe*, *55*(2), 289-303.

Read Write Inc. Fresh Start

Gorard, S., Siddiqui, N., & See, B. H. (2016). An evaluation of Fresh Start as a catch-up intervention: a trial conducted by teachers. *Educational Studies*, *42*(1), 98-113.

SAFMEDS

Beverley, M., Hughes, J. C., & Hastings, R. P. (2016). Using SAFMEDS to assist language learners to acquire second-language vocabulary. *European Journal of Behavior Analysis*, *17*(2), 131-141

Byrnes ME, Macfarlane CA, Young KR, West RP. (1990) Using precision teaching to teach minimum competency test skills. *Teaching Exceptional Children.* 22:58–61. doi:10.1177/004005999002200415.

Casey, J., McLaughlin, T. F., Weber, K. P., & Everson, M. (2003). The effects of five minute practice, unlimited practice, with SAFMED cards on correct and error rate in math facts for two elementary school children with learning disabilities. *International Journal of Special Education*, *18*(1), 66-72.

Cunningham, D., McLaughlin, T. F., & Weber, K. P. (2012). The use and evaluation of verbal prompting with see to say problems and answers and SAFMEDS to teach math facts to a student with learning disabilities. *International Journal of Advances in Psychology*, *1*(2), 37-39.

Greene, I., Mc Tiernan, A., & Holloway, J. (2018). Cross-age peer tutoring and fluency-based instruction to achieve fluency with mathematics computation skills: A randomized controlled trial. *Journal of Behavioral Education*, *27*(2), 145-171.

Hughes, J. C., Beverley, M., & Whitehead, J. (2007). Using precision teaching to increase the fluency of word reading with problem readers. *European Journal of Behavior Analysis*, *8*(2), 221-238

Hunter, S. H., Beverley, M., Parkinson, J., & Hughes, J. C. (2016). Increasing high school students' maths skills with the use of SAFMEDS class-wide. *European Journal of Behavior Analysis*, *17*(2), 154-165.

Lambe, D., Murphy, C., & Kelly, M. E. (2015). The impact of a precision teaching intervention on the reading fluency of typically developing children. *Behavioral Interventions*, *30*(4), 364-377.

Quigley, S. P., Peterson, S. M., Frieder, J. E., & Peck, K. M. (2018). A review of SAFMEDS: evidence for procedures, outcomes and directions for future research. *Perspectives on Behavior Science*, *41*(1), 283-301

Shue, S. (2017). *Vocabulary acquisition and behavior outcomes for students with EBD: Flashcard usage and instructional delivery* (Doctoral dissertation, University of Pittsburgh).

Seasons for Growth

Riley, A. (2012). Exploring the effects of the 'Seasons for Growth' intervention for pupils experiencing change and loss. *Educational and Child Psychology*, *29*(3), 38.

Social Stories

Beh-Pajooh, A., Ahmadi, A., Shokoohi-Yekta, M., & Asgary, A. (2011). The effect of social stories on reduction of challenging behaviours in autistic children. *Procedia-Social and Behavioral Sciences*, *15*, 351-355.

Chan, J. M., & O'Reilly, M. F. (2008). A Social Stories™ intervention package for students with autism in inclusive classroom settings. *Journal of Applied Behavior Analysis*, *41*(3), 405-409.

Chan, J. M., O'Reilly, M. F., Lang, R. B., Boutot, E. A., White, P. J., Pierce, N., & Baker, S. (2011). Evaluation of a Social Stories™ intervention implemented by pre-service teachers for students with autism in general education settings. *Research in Autism Spectrum Disorders*, *5*(2), 715-721.

Da Silva, M. C., Arantes, A., & Elias, N. C. (2020). Use of social stories in classroom to children with autism. Psicologia Em Estudo, 25, 15. Retrieved from https://www-proquest-com.ezproxy.bangor.ac.uk/scholarly-journals/use-social-stories-classroom-children-with-autism/docview/2393665621/se-2?accountid=14874

Delano, M., & Snell, M. E. (2006). The effects of social stories on the social engagement of children with autism. *Journal of positive behavior interventions*, *8*(1), 29-42.

Hanrahan, R., Smith, E., Johnson, H., Constantin, A., & Brosnan, M. (2020). A pilot randomised control trial of digitally-mediated social stories for children on the autism spectrum. *Journal of Autism and Developmental Disorders*, *50*(12), 4243-4257.

Kim, M. S., Blair, K. S. C., & Lim, K. W. (2014). Using tablet assisted Social Stories™ to improve classroom behavior for adolescents with intellectual disabilities. *Research in Developmental Disabilities*, *35*(9), 2241-2251.

Malmberg, D. B., Charlop, M. H., & Gershfeld, S. J. (2015). A two experiment treatment comparison study: Teaching social skills to children with autism spectrum disorder. *Journal of Developmental and Physical Disabilities*, *27*(3), 375-392.

O'Connor, K. M., & Hayes, B. (2019). A real-world application of Social Stories as an intervention for children with communication and behaviour difficulties. *Emotional and Behavioural Difficulties*, *24*(4), 323-338.

Saad, M. A. E. (2016). The Effectiveness of Social Stories among Children and Adolescents with Autism Spectrum Disorders: Meta-Analysis. *International Journal of Psycho-Educational Sciences*, 5(2), 10.

Sansosti, F. J., & Powell-Smith, K. A. (2006). Using social stories to improve the social behavior of children with Asperger syndrome. *Journal of Positive Behavior Interventions*, *8*(1), 43-57.

Schneider, N., & Goldstein, H. (2009). Social Stories™ improve the on-task behavior of children with language impairment. *Journal of Early Intervention*, *31*(3), 250-264.

Silva, M. C. D., Arantes, A., & Elias, N. C. (2020). Use of social stories in classroom to children with autism 1. *Psicologia em Estudo, 25*.

Srija, S. S., Sugi, S., & Srinivasan, D. (2019). Efficacy of social stories in coping with bullying among children with childhood psychiatry conditions: A randomized controlled interventional study. *The Indian Journal of Occupational Therapy*, *51*(2), 52.

Thompson, R. M., & Johnston, S. (2013). Use of social stories to improve self-regulation in children with autism spectrum disorders. *Physical & Occupational Therapy in Pediatrics*, *33*(3), 271-284.

Student Assistance Program (SAP)

Biddle, V. S., Kern III, J., Brent, D. A., Thurkettle, M. A., Puskar, K. R., & Sekula, L. K. (2014). Student assistance program outcomes for students at risk for suicide. *The Journal of School Nursing*, *30*(3), 173-186.

Time-out

Donaldson, J. M., & Vollmer, T. R. (2011). An evaluation and comparison of time-out procedures with and without release contingencies. *Journal of Applied Behavior Analysis*, *44*(4), 693-705.

Vegas, K. C., Jenson, W. R., & Kircher, J. C. (2007). A single-subject meta-analysis of the effectiveness of time-out in reducing disruptive classroom behavior. *Behavioral Disorders*, 32(2), 109-121. Retrieved from https://www-proquest-com.ezproxy.bangor.ac.uk/scholarly-journals/single-subject-meta-analysis-effectiveness-time/docview/219678709/se-2?accountid=14874

Wolf, T. L., McLaughlin, T. F., & Williams, R. L. (2006). Time-Out Interventions and Strategies: A Brief Review and Recommendations. *International Journal of Special Education*, *21*(3), 22-29.

Toe by Toe

Jeffes, B. (2016). Raising the reading skills of secondary-age students with severe and persistent reading difficulties: Evaluation of the efficacy and implementation of a phonics-based intervention programme. *Educational Psychology in Practice*, 32(1), 73-84.

O'Rourke, D., Olshtroon, A., & O'Halloran, C. (2016). The Limerick Reading Initiative: a reading intervention targeted at struggling readers in primary school. *Support for Learning*, 31(2), 148-163.

Chapter 4: The factors that inform school leaders’ choice of teaching programmes

Pegram, J., Watkins, R., Hoerger, M., & Hughes, J. C (under review). The factors that inform school leaders’ choice of teaching programmes. Manuscript under review.

Abstract

Evidence suggests that some teachers are now engaging with the external research evidence. However, engagement does not necessarily translate into evidence-based teaching and learning, and improved outcomes. When making decisions about which teaching programmes and practices to adopt, recent evidence from survey research suggests that teachers' decisions are greatly influenced by their own experiences, the experiences of colleagues and staff in other schools, and by non-research-based professional development. The adoption process is complex and a better understanding of this may provide insight into the best way to increase the uptake of evidence-based programmes and practices. We conducted this study with a cluster of schools in Wales to explore in greater depth the factors that influence head teachers' decisions when choosing which teaching programmes to adopt. We conducted interviews with the head teachers from two secondary schools, six primary schools, and one special school. Our results support previous survey findings and provide greater knowledge of the interrelated factors that influence decisions to adopt programmes. We address gaps in the knowledge and suggest how policymakers, researchers, and school improvement professionals can improve the transmission and adoption of evidence-informed teaching practices.

Over the last 20 years there has been an increasing emphasis on the value of evidence-based education to raise standards in schools (e.g., Cooper et al., 2009; Department for Education, 2010, 2016a; Institute for Effective Education, 2019; OECD, 2017; Welsh Government, 2015). There is growing evidence to suggest that access to, and use of, research findings have a positive impact on the knowledge, skills and confidence of educators, and positively impacts teaching provision (Brown & Greany, 2018; Cordingley et al., 2015; CUREE, 2010; Greany, 2015; Godfrey, 2017; Mincu, 2015; Nelson & O'Beirne, 2014). Through using research and evidence to inform decision making, school leaders are more likely to spend scarce resources, such as time and money, more efficiently to improve the quality of provision (Gorard, See, & Siddiqui, 2020; Oxman et al., 2009; Styles & Torgerson, 2018). However, despite the focus on the use of evidence and research in education, there is currently very little evidence to suggest that research use directly impacts learner outcomes (Godfrey & Brown 2018; Gorard, See & Siddiqui, 2020).

There are two distinct forms of research evidence that can be used to improve teaching and learning. First, evidence derived from the external research community including peer-reviewed and published primary studies and reviews, typically produced by academic researchers. Second, practitioner-led evidence that is derived through action research and the collation of school and/or system level data (Brown & Greany, 2018). Over recent years, a number of research-engaged schools have been established in the United Kingdom, generating their own practitioner-led evidence and drawing on evidence from the wider research community to improve practice and, ultimately, learner outcomes (e.g., Godfrey, 2016; Nelson & Campbell, 2017; Sharples, 2013).

There remains an ongoing debate about what constitutes reliable research evidence in education. Many policymakers and scholars favour evidence from high-quality experimental research, which includes the use of randomised controlled trials or other quasi-experimental designs (Goldacre, 2013). The creation of research organisations such as the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), the Research School Network run by the Education Endowment Foundation (EEF), Evidence 4 impact (E4I) and the Best Evidence Encyclopaedia (BEE) reflects a wider movement – both in government and research circles - towards the promotion of a more evidence-based education system (Welsh Government, 2021). However, there are concerns about the problems associated with a ‘prescriptive’ approach to education policymaking, which some argue does not take into account the multiple and complex factors that influence decision-making processes when adopting programmes and practices in schools (such as Biesta, 2007; Bristow et al., 2015; Nutley et al., 2013; Sharples, 2013). In Wales, recent government policy initiatives have outlined a commitment to developing an evidence-informed approach to education policy and practice, which includes the use of the best available international research evidence and evidence from teacher professional enquiry (Welsh Government, 2021).

There are a host of theoretical frameworks that attempt to describe the process of implementing new ideas and practices (implementation of innovations) within an organisational context, and several reviews have attempted to synthesise the broad range of literature (e.g., Dooley, 1999; Greenhalgh et al., 2004; Kapoor, Dwivedi, & Williams, 2014; Wisdom et al., 2014). One of the earliest ideas, developed by Rogers (1995), was the diffusion of innovations theory, which defines how new ideas or practices spread (diffuse) through a social system. Rogers’ (2003) diffusion theory comprises five main phases, with the first three phases defining the pre-implementation or adoption phase relevant to the scope of this study. Phase 1 is *knowledge* – an adopter becomes aware of an innovation and how it functions. For example, an adopter will seek an innovation that is compatible with existing values, needs and resources; one that is cost-effective, impactful, simple to use and adaptable, and can be trialled and the perceived benefits observed. Phase 2 is *persuasion* – the adopter forms an attitude towards the innovation, initiated by a change agent, which may be inter-organisational or external to the organisation but is more effective if their contextual backgrounds are similar. Phase 3 is *decision* – the adopter engages in a decision to adopt or reject the innovation. More recent evidence from a narrative synthesis of adoption of innovation theories by Wisdom et al. (2014) reports similar findings. For example, inter-organisational and external social networks influence decisions to adopt, as do the characteristics of an innovation. Innovations that are clear and simple to use, advantageous (cost effective, feasible and beneficial), require minimal resources to implement, observable, and compatible with existing norms and values are critical facilitators of the adoption process. Wisdom et al. also found that innovations deemed compatible with existing practice and with clear research evidence and practice efficacy, are more likely to be considered.

To date, there remains a lack of consistent and regular use of external research evidence by teachers to inform teaching practice (Brown & Zhang, 2016; Coldwell et al., 2017; Nelson et al., 2017; Walker et al., 2019). Historically, educators have been unable to access the external research and have lacked the necessary skills to understand it (Hemsley-Brown and Sharp, 2003; Schaik et al., 2018). Whilst these barriers still exist, the creation of research repositories has allowed some evidence to be available in a more accessible format. However, the results from two teacher surveys found that the availability of evidence repositories such as the EEF only had a small to moderate influence on decision-making in schools (Nelson et al., 2017; Walker et al., 2019).

Despite the greater availability of research evidence, one of the barriers to the uptake of evidence in schools is the lack of relevant, practical, and useable information (Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018). Results from a systematic review undertaken by Connelly et al. (2018) indicate that all randomised controlled trials conducted in education from 1980 to 2016 neglected to address implementation and causal process. The consequence of this is an evidence-base that does not provide teachers with the information required to understand if an intervention programme or strategy has utility in their school setting. Furthermore, the proliferation of research literature, research repositories, and differing criteria for rating the strength of evidence, has resulted in a growing evidence base that is viewed by many as disjointed, complex and often contradictory and inconclusive. Importantly, this can result in teachers feeling sceptical about the value of education research (Broekkamp & Van Hout-Wolters, 2007; Gorard et al., 2020).

To encourage the use of research and evidence in schools, it is suggested that school leaders should promote a culture of evidence use and demonstrate their alignment to an evidence-informed teaching approach (Brown & Zhang, 2016; Coldwell et al., 2017; Nelson & O'Beirne, 2014). According to an attitude survey by Ager and Pyle (2013), 67 per cent of head teachers claim to have consulted the evidence base when planning school provision, with nearly half referring to the EEF Toolkit when deciding how to spend their Pupil Premium (Ager & Pyle, 2013). A more recent survey published by the Sutton Trust (2018), found that 59 per cent of senior leaders now access the Teaching and Learning Toolkit to inform approaches to teaching and learning. Whilst this presents some evidence of engagement, it does not determine the extent that decision-making is underpinned by research evidence (Higgins, 2016). Gorard et al. (2020) proposes that evidence from external sources may increase user knowledge but has very little impact on teaching practice. Results from a study by Pegram et al. (2022 – Chapter 3) found that 30 per cent of all interventions in use across ten schools had causal evidence of impact on pupil outcomes, 67 per cent had no evidence, and three per cent had evidence of ineffectiveness. These results suggest that an increased awareness of the use of research and evidence across education settings over recent years has not necessarily translated into evidence-based decision making in schools.

If, as previously suggested, the evidence-base lacks the information that teachers require when decision making, then it is unsurprising that teachers more often turn to the expertise of colleagues as the best available source of information (Walker et al., 2019). Recent research indicates that teachers more often draw on their own experiences and the expertise and views of colleagues and staff in other schools when choosing which programmes and practices to adopt (Nelson et al., 2017; Nelson & O'Beirne, 2014; Walker et al., 2019). In addition, Nelson et al. (2017) and Walker et al. (2019) found that non-research-based professional development events also had a major influence on approaches to teaching and learning.

Much of the research conducted on this topic has primarily aimed to examine teachers' attitudes and engagement with research and evidence (Brown & Zhang, 2016; Coldwell et al., 2017; EEF, 2018; Nelson & O'Beirne, 2014). Other researchers have used surveys to explore the other factors that inform teachers' decisions when choosing which programmes or practices to adopt (Nelson et al, 2017; Walker et al., 2019). Whilst the findings from these surveys are useful, there is currently a lack of research that evaluates in more detail the factors that influence these decision-making processes in schools. We were interested to explore the process by which schools identify which programmes and interventions to use as this will provide an important insight into the best way to increase the uptake of evidence-based programmes and practices.

We conducted the following study with an established cluster of schools in Wales to explore the factors that inform school leaders' decisions when choosing which programmes and practices to adopt. The cluster included two secondary schools (English medium), seven primary schools (one Welsh medium and six English medium) and one special school (English medium), located in an area of high deprivation. With the intention of moving towards a more evidence-informed practice, the cluster had entered into a research partnership with the Regional School Effectiveness and Improvement Service for North Wales (GwE) and the Collaborative Institute for Education Research, Evidence and Impact (CIEREI), Bangor University, to commission this evaluation work alongside other close-to-practice research projects. Throughout the project, the head teachers and senior leaders were closely involved in the development and planning of research and worked collaboratively with the lead researcher who was embedded within the cluster.

Through semi-structured interviews, we aimed to explore the factors that inform decision-making when choosing school provision. We asked the following research question: What factors influence head teachers' decision-making when choosing which intervention programmes to adopt?

Method

Ethics

The study was conducted under ethical approval from the School of Education Research Ethics Committee, at Bangor University (ref: 18-03). Consent was obtained from the participating head teacher in each school. Participants were provided with examples of the questions before commencing the interviews. To protect anonymity, identifiable information such as school and head teacher name are not included in this paper.

Participants

We conducted our study with nine head teachers from a cluster of ten schools made up of seven primary schools (six English medium and one Welsh medium), two secondary schools (English medium), and a special needs school (English medium). The schools are in an area of high deprivation in a large town in Wales. The first author was unable to organise an interview with the remaining head teacher.

Interview design and procedure

When designing the semi-structured interview questions, we drew on the principles used in the survey studies conducted by Nelson et al. (2017) and Walker et al. (2017), and chose to focus the interview questions to a specific programme or strategy that the head teacher had recently adopted. This meant we could gain a realistic insight into the variety of factors that inform adoption-decisions in the school. We asked the head teachers two very similar interview questions: to think of a literacy or numeracy programme that they recently adopted and explain why they chose it (question 1); and, to think of a well-being programme that they recently adopted and explain why they chose it (question 2). We chose these two questions based on a preconception that there may be differences in decision making for the two broad areas of provision (academic and well-being provision). To extract more detail, we prepared the following three probes: What factors influenced your decision? Did you consult anyone? Did you access information? For the purpose of this study, the participants' characteristics were not deemed relevant.

An email invite was sent to each of the head teachers inviting them to attend a one-to-one recorded interview with the first author. The interviews were conducted during the summer term of 2018 in a meeting room in each school. Each interview lasted around 30 minutes.

Analysis

We chose to adopt the Reflexive Thematic Approach (RTA) outlined by Braun and Clarke (2013, 2019, 2020). This approach was chosen because it acknowledges the impact of the researcher on the analytic process and on knowledge creation. Unlike other more structured approaches to thematic analysis, the RTA process requires an organic and recursive coding approach, and advocates for deep reflection on, and engagement with, data (Braun & Clarke, 2019). We adopted a constructionist epistemology, and a predominantly inductive approach to evaluate the data most effectively. However, a degree of deductive analysis was employed to ensure that the coding contributed to producing themes meaningful to the research questions. Furthermore, previous survey research indicated a number of likely themes that might emerge during the course of this study (e.g., Nelson et al, 2017; Walker et al., 2019).

We analysed the data using the six stages of RTA as outlined by Braun and Clarke (2013; 2019; 2021). In phase 1 we transcribed the recorded data and read the dataset twice to establish familiarity; phase 2 involved the generation of initial codes for the salient and interesting comments; in phase 3 the coded data was reviewed and analysed to generate themes; in phase 4 a recursive review of potential themes was conducted and themes were refined; in phase 5 the themes were defined and named; and, in phase 6 a narrative description was created using the themes. In accordance with the reflexive approach of thematic analysis, we refrained from pursuing consensus from an independent reviewer/coder. RTA is “the researcher’s reflective and thoughtful engagement with their data and their reflexive and thoughtful engagement with the analytic process” (Braun and Clarke 2019, p. 594). Figure 1 and 2 displays the themes derived from the data set and the relationships between themes.

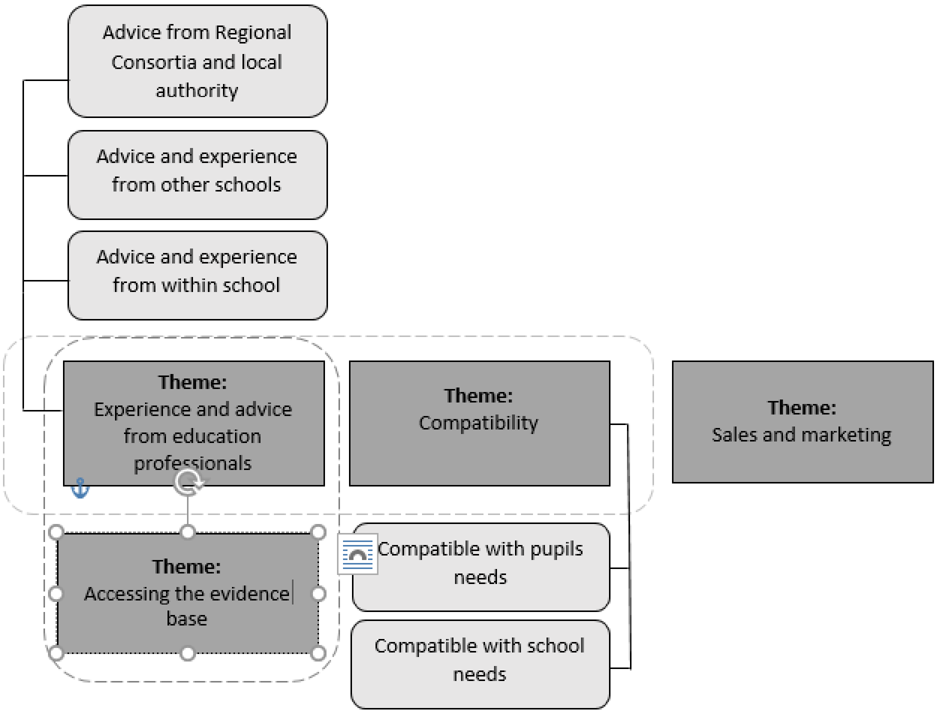

Advice from Regional Consortia and local authority
Advice and experience from other schools
Advice and experience from within school
Theme: Experience and advice from education professionals
Theme: Compatibility
Theme: Sales and marketing
Theme: Accessing the evidence base
Compatible with pupils needs
Compatible with school needs

Figure 1. Themes derived from the data for question 1 (think of a literacy or numeracy programme that you recently adopted and explain why you chose it)

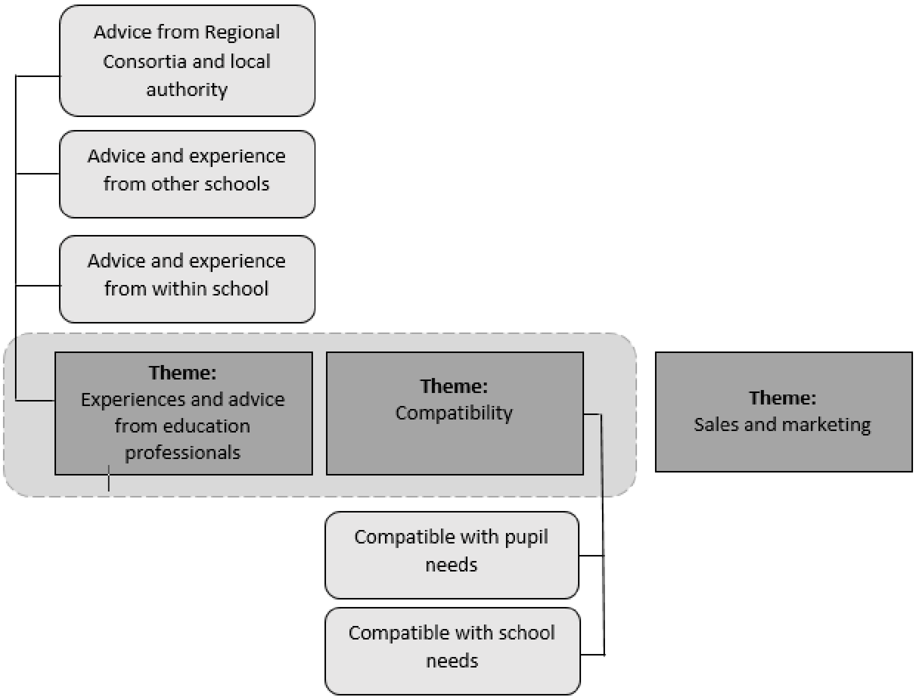

Advice from Regional Consortia and local authority
Advice and experience from other schools
Advice and experience from within school
Theme: Experiences and advice from education professionals
Theme: Compatibility
Theme: Sales and marketing
Compatible with pupil needs
Compatible with school needs

Figure 2. Themes derived from the data for question 2 (think of a well-being programme that you recently adopted and explain why you chose it)

Results

Theme: Experience and advice from education professionals

This theme and the associated three subthemes were identified from discussions about both academic and well-being programmes.

Advice and experience from within school

Several of the head teachers referred to drawing on their own experience when selecting programmes or strategies. As described in extract 1 below, the head teacher chose to use Catch-Up reading because they were already familiar with the functions of the programme. Extract 2 illustrates how one of the head teachers used their own skills and experience to create a bespoke approach to mindfulness. A common occurrence throughout the data set was that the head teachers would seek advice and experience from colleagues within school when choosing both academic and well-being programmes. Head teachers indicated that when seeking advice internally, they would more often consult with members of the senior leadership team (extract 3). Another head teacher described how the school encourages a process of collaborative decision-making between teachers, the head teacher and senior leadership team when adopting new ideas (extract 4).

> Extract 1: “The academic programme that I've implemented since September is Catch-Up reading, so I chose Catch-Up reading because I was familiar with this”.
>
> Extract 2: “I've implemented our own approach to mindfulness since I've been here. So all staff have been trained, parents been trained, we're doing class by class delivering mindfulness sessions. But it's not something I've purchased, we put it together ourselves… It was my own background, I came from a school where actually we developed a mindfulness curriculum, and it's also where I would like us to go.”
>
> Extract 3: “I meet with my senior leadership team and my well-being mentor, and we discuss things”.
>
> Extract 4: “I think it's the culture within the school it's not just from me, if anybody has an idea, they write a paper. Just a couple of ideas about how much it's going to cost and what they think the impact is going to be. They write a paper, email it to me, I read it and discuss it in the senior management meeting”.

Advice and experience from other schools

When choosing academic and well-being programmes, all of the head teachers referred to seeking advice from and drawing on the direct experiences of head teachers already using the programmes in other schools. As illustrated in extract 1, the head teacher sought other schools using the Place2be programme (a service designed to improve mental health), so they could observe evidence of effectiveness and understand the causal processes and functions of the programme that promote or hinder effectiveness. This helped the head teacher to determine the compatibility of the programme and the likely impact in their context. Extract 2 provides a similar example of how a head teacher adopted a programme based on an observable experience in another school. It was evident from the comments that the head teachers valued and trusted the advice from colleagues in other schools, particularly those that were homophilous (for example, extract 3 presents a good example of this). All of the head teachers from the cluster that participated in the study referred to consulting with and drawing on the advice of other head teachers from within the cluster, and extract 4 provides a good illustrative example of a how a programme could be adopted or rejected based on advice from within the cluster.

> Extract 1: “I would contact people that I know in the cluster but also out of the cluster and across County that I know have done similar work. We are going to have a look at Place2be after Easter, because we would like to see it in action. Because I think if you actually go to a school where it's working, they can tell you all the pitfalls as well as the benefits, and guide you through it.”
>
> Extract 2: “We first saw it in a meeting over in the xxx cluster, they were trying it, and it just seemed to hit the nail on the head for what some of our children were struggling with.”
>
> Extract 3: “We consult other professionals and go to other schools and talk to people we can see have lived our life, walked in our shoes and delivered in spite of it. They are the people we trust!”
>
> Extract 4: “Certainly all the heads in the cluster. We would discuss it in our regular meetings, and if anybody said I found this programme, we’ve trialled it, it is working. You know I think it is good. Or we found this program and actually it was a waste of time.”

Advice from regional consortia and local authority

Many of the head teachers’ comments referred to seeking out advice and/or following recommendations from education professionals from their external network. Across the data set it was common for the head teachers to refer to recommendations from the regional consortia school improvement advisor. Extract 1 presents an example of how one school adopted a reading fluency programme after using it as part of a research project, which was promoted by the school improvement advisor. In extract 2, the head teacher indicated that they actively seek advice about evidence-based practices from the regional school improvement advisor when considering well-being provision. Moreover, the head teacher refers to viewing the role of the regional school improvement advisor as a facilitator of evidence-based practice. Again, in extract 3, the head teacher identified that the regional school improvement advisor provides recommendations about best practice methods that are compatible with the needs of the school. When considering which well-being programmes to adopt, the head teachers often draw on the advice and recommendations from education professionals in local authority such as the healthy schools well-being co-ordinator, welfare officer and educational psychologist (for example, see extracts 4 and 5 for evidence of this). Furthermore, in extract 4 the head teacher referred to the healthy schools co-ordinator as being knowledgeable about the needs of the school.

> Extract 1: “Headsprout came about as part of a research project, a joint working project with xxx xxx (supporting improvement adviser from the regional school improvement service) a couple of years ago, with a group of schools. I had never heard of it before then so that's how it came about.”
>
> Extract 2: “Xxx (supporting improvement adviser from the regional school improvement service) would be the first port of call. Years ago I would have probably done it and said it is my gut reaction, I'm doing this. Whereas now, what is the research saying? If there is research can we find out what it's saying about it. We're doing some work at the moment on all sorts of things, well-being is one of them. PSE through the school. Where is the gaps in it? What does the research say to improve it? And that is their (supporting improvement adviser from the regional school improvement service) job in the cluster to find out more about it.”
>
> Extract 3: “We have regular visits from xxx (supporting improvement adviser from the regional school improvement service) to monitor standards and things like that. When working together, sometimes areas are identified that need a bit of work on and because they have got so many contacts in other schools and are aware of initiatives, they might guide to something that might suit our school.”
>
> Extract 4: “We have a well-being member of staff who is healthy schools and well-being coordinator, and she works closely with xxx xxx, who is the deputy county council healthy schools well-being person, who makes schools coordinators aware of initiatives and training that are going on, which schools might be interested in. She is a regular visitor to the school. She knows some of the issues that we deal with.”
>
> Extract 5: “I would ask the education welfare officer and the educational psychologists. I would definitely because we are looking at putting PASS into our school, and want to know what would be appropriate, and for what groups. I would definitely ask the educational psychologist.”

Theme: Compatibility

This theme and subthemes were identified from discussions about both academic and well-being programmes.

Compatible with pupil needs

It was clear from many of the comments that the needs of pupils often primarily influenced decisions about which programmes to adopt. Several of the head teachers referred to choosing a programme based on its functional compatibility with existing pupil needs. For example, in extracts 1 and 2 the head teachers describe how the functions of the programmes were compatible with the literacy and mental health needs of pupils. Similarly, in extracts 3 and 4, the head teachers revealed that the primary reason for choosing the programmes was to support the pupils with learning and behavioural difficulties.

> Extract 1: “To fit in with what we want to do. We’ve always got children in school that are struggling with reading and comprehension throughout the school”… “This one seemed to hit the reading early on, and particular the comprehension element of it is good, drawing on the skills that they need”.
>
> Extract 2: “It's very topical at the moment, mental health, the lack of mental health facilities are everywhere, and especially for vulnerable pupils and their parents, it's even more so. So it's about trying to give them strategies they can independently apply”.
>
> Extract 3: “The reason we chose it was because children were falling behind in their literacy skills. And having conducted an audit it was pretty apparent that some children were behind with their literacy skills. We felt that we needed to find something that would make a difference to those children that are falling behind and that's how we ended up with Catch-Up Literacy. Catch-Up numeracy came on after literacy for the same reason really we could see that the children needed a boost with basic skills to achieve their full potential”.
>
> Extract 4: “It was the year group that had a high number of behaviour and learning difficulties; and that's what prompted that”.

Compatible with school needs

The importance placed on the compatibility of programmes with the needs of the school was evident throughout the dataset. Most of the head teachers talked about characteristics of the programmes that suited the needs of the school. They often referred to choosing programmes based on their efficiency and ability to have minimal impact on staff time. For example, in extract 1 the head teacher refers to choosing programmes that require minimal support from one or two staff members, and highlights that programmes delivered via digital platforms are important. In extract 2, the head teacher refers to the ability to conduct with multiple learners at differing levels. Moreover, in extract 3, the online functionality and ability to use at home, prompted its adoption.

Several of the head teachers described adopting programmes because they were easy to use and clear. In extract 4 the head teacher describes the importance of using programmes or approaches that are simple to use and require limited preparation. In this example, the ability to link with a mobile phone application was also important. Again, in extract 5, the head teacher highlights the need for simplicity, ease of use and availability of resources. In extract 6, the head teacher illustrates how a programme requires only minimal resources, is efficient and easy to deliver.

> Extract 1: “We look for a programme that can be done alongside an adult maybe two children doing it or an adult working with someone else whilst keeping an eye. So having something computerised is quite important to us”.
>
> Extract 2: “A lot of children can do it at once, so time-wise, you can have a room full of children at different levels”.
>
> Extract 3: “There's nothing they can access at home, so we've now bought into Sumdog, which is gaming, but maths gaming online. So were using that and again we're seeing that children are accessing that an awful lot at home”.
>
> Extract 4: “We’ve tried to keep it as simple as possible, there is nothing complicated, there is not a massive amount of preparation. It is something that anybody could pick up and go. It is something that works really well with IT and the mindfulness apps, so we tried to combine it, so you can have it on your mobile phone. Really keep it simple”.
>
> Extract 5: “Again, I think going back to the whole package, they have the whole pre and post assessments with it, it was very structured, it was scripted, and even all the resources were in it”.
>
> Extract 6: “It's quite easy to deliver and the children can get on and do it themselves. The time allocation is not too draining on the staff and the timetable”.

Theme: Sales and marketing

This theme was identified from discussions about both academic and well-being programmes.

Several of the head teachers' comments, relevant to this theme, described adopting programmes after attending a course or seeing it on a website (see extract 1 and 2). Two of the head teachers specifically referred to programmes being sold well at the training course (see extract 3 and 4). One head teacher referred to adopting a programme based on the marketing on the website.

> Extract 1: "It was at a course I went on, based on speech and language. Talk Boost!"
>
> Extract 2: "The reason we picked IDL, the SENCO had been on training, and had seen it there".
>
> Extract 3: "I chose that because I went on a training course for it and to be honest it was sold very well on the training course... They were talking about places that it had been used, and the benefits, and other places nationwide. They did mention a bit of research."
>
> Extract 4: "Because it looked from the website a very good programme".

Theme: Accessing the evidence-base

This theme was only identified from discussions about academic programmes.

Two of the head teachers interviewed referred to adopting programmes based on an established evidence base. See extract 1 and 2. In extract 2, the head teacher also refers to having previous knowledge of Catch-Up reading, which equally influenced the decision.

> Extract 1: "It's well-researched. We trained up the TA's in order to deliver it, it's been delivered across the school consistently. It's got a good body of knowledge behind it".
>
> Extract 2: "So I chose Catch up reading because I was familiar with it and it was well-researched".

Discussion

Since the creation of the education research repositories such as the EEF Toolkit (Higgins et al., 2016), more schools now access the external research evidence to inform their provision (Sutton Trust, 2018). However, when choosing which programmes to adopt it seems that decision-making is rarely influenced by the external research evidence (Walker et al., 2019). Most of the existing survey-based research focuses primarily on teachers' engagement with research evidence, which often neglects to consider context (Nelson et al., 2017; Walker et al., 2019). Furthermore, the nature of closed response surveys means that there is currently a paucity of research that evaluates in more detail the range of factors that influence the decision-making processes in schools. Through the use of semi-structured interviews, this study aimed to gain a more complete understanding of the factors that influence head teachers' decisions when choosing which academic and well-being programmes to adopt. The adoption process is complex, and we believe that a better understanding of how head teachers choose programmes and interventions may help us improve the uptake of evidence-based programmes and practices in schools.

We asked the head teachers two similar interview questions related to specific programmes they had recently adopted. The first asked head teachers to explain why they chose an academic programme, and the second asked them to explain why they adopted a well-being programme. We analysed the data using thematic analysis and this revealed a total of four main themes across the dataset. Notably, three of the themes (experience and advice from education professionals, compatibility, and sales and marketing) were derived from the data when discussing both academic and well-being programmes, and the fourth theme (accessing the evidence base) was only found in the data relevant to academic programmes. As expected, most of the themes were interrelated. For example, experience and advice from education professionals was interrelated with compatibility, and experience and advice from education professionals was linked to accessing the evidence base.

Data from recent survey research conducted by Nelson et al. (2017) and Walker et al. (2019) identified that teachers were more likely to seek ideas and support from their own experiences, and from other teachers and schools, when choosing approaches to teaching and learning. Our results support these findings and expand them. Not only did we find that the head teachers referred to drawing on their own experience, and the experience and advice from colleagues within school, we found that head teachers would most often consult with and draw on the experience of other head teachers within the school cluster and in other schools across the region, and occasionally beyond. According to Rogers' (1995) diffusion of innovation theory, adopters will form opinions based on the ideas of others within their social networks, particularly if contextual backgrounds are similar. Importantly, we discovered that it was common for a head teacher to visit another school to observe a programme of interest in use; not only to determine compatibility, but also to understand the characteristics of the programme that promote or hinder success. Previous research suggests that inter-organisational relationships influence decisions to adopt innovations (Wisdom et al., 2014). Our data also suggests that head teachers trust and value the advice and experience from other schools, and will often adopt or reject a programme based on the information they receive.

Our data also supports findings from Rogers (1995) and Wisdom et al. (2014) and indicates that advice is also frequently sought from external social networks. More specifically, the head teachers would solicit advice from their linked regional school improvement adviser, an educational psychologist and the healthy schools' co-ordinator. We discovered that the head teachers more often referred to seeking out and/or following recommendations from their school improvement adviser (from the regional school improvement service), who was a trusted source of knowledge for research and evidence-based practices.

Previous research provides evidence to suggest that adopters will search for information about the characteristics or functions of an innovation to establish its compatibility with existing values and needs (Rogers, 1995). The head teachers we interviewed highlighted the influence that pupil mental health and academic needs had on the adoption of programmes. Some of the comments suggested that the needs of particular groups of pupils often led the head teachers to seek out programmes that were functionally compatible. Wisdom et al. (2014) identified that innovations viewed as simple to use, advantageous, observable, and of minimal impact on resources are some of the critical factors that facilitate adoption. Similarly, the participants in our study often referred to choosing programmes on the following key features: they were simple to use; placed minimal demand on resources (staff time and money); were time saving; they could be delivered by computers or digital platforms; and could be accessed at home.

The evidence from previous survey research by Walker et al. (2019) found that teachers' decisions were also influenced by continuing non-research-based professional development events, and our findings offer some support for this. For example, a few head teachers referred to adopting programmes after attending a course. Our data also revealed that some head teachers adopted programmes following an effective sales pitch at a training course. Finally, one head teacher described how a programme was marketed effectively on a website and this led to the adoption of the resource.

Significantly, only two of the head teachers referred to choosing academic programmes as a result of an established evidence-base (research evidence from external sources). In one of the comments the participant also refers to being influenced by previous experience of using the programme. Notably, there was no mention of research evidence influencing decisions when choosing well-being programmes. Therefore, the findings from this study might suggest that the external evidence-base still lacks relevant, practical and useable summaries of evidence (Connelly et al., 2018; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018).

Whilst we have been able to collect rich contextual data from a cluster of schools in Wales, we recognise that data derived from a small number of participants may not necessarily generalise to other schools. However, our findings support the results from recent questionnaire-based studies (Nelson et al 2017; Walker et al., 2019) and align very well with many of the constructs from existing adoption of innovation theories (Rogers, 2003; Wisdom et al., 2014). It is important that we also draw attention to the impact that existing knowledge, experience and opinions of the first author may have had on the results, particularly as the first author was embedded within the cluster as part of a three year research studentship.

The results from this study provide a greater depth of understanding of the interrelated factors that influence head teachers' decisions when choosing school provision. Gaining a better understanding of these factors, provides policy makers and researchers with information that helps to determine the best way to increase the adoption of evidence-based teaching programmes and practices. Future research might consider a replication of this study, with senior leaders and teachers from other school clusters across the UK. Further research could also be directed at answering the following questions. First, how can policy makers ensure that factors such as the compatibility and usability of evidence-based programmes and practices are communicated through school networks? Second, how can we ensure that information about evidence-based programmes and practices are communicated to school networks more effectively, and how do we prevent the spread of non-evidence-based and ineffective programmes and practices?

Our findings suggest that head teachers require information to determine which programmes and practices are likely to be compatible with culture, needs and available resources, and the evidence-base currently lacks this type of relevant and usable evidence (Connelly et al., 2018; Van Schaik et al., 2018). If schools are to adopt provision that has positive causal evidence, then we suggest that the evidence-base must also include evidence from process evaluations and social validity research together with important information on effective implementation (EEF, 2018, 2019). Consequently, it is imperative that researchers provide practitioners with a more relevant and practical evidence-base to help determine what has the greatest chance of working in what context. This is an important role for researchers to improve the quality of schools' decision-making during the *knowledge* and *persuasion* phase of pre-implementation. We also know that the adoption of new programmes and ideas is influenced by existing relationships within inter-organisational and external social networks. For that reason, we suggest that external relationships be nurtured, and a more formalised and planned process for disseminating evidence-based programmes and practices be developed using change agents. According to Nelson et al. (2017) and Coldwell et al. (2017), teachers are more likely to engage in evidence-based practice if it is supported by evidence from 'trusted practitioners' and/or a 'trusted conduit' such as educational charities or knowledge brokers (Gorard et al., 2020). To improve the quality of decision-making in schools at the *knowledge* and *persuasion* phases of the pre-implementation process, we must improve the transmission of evidence-based programmes and practices between schools. We propose the following two practical solutions. First, create more opportunities for researchers and school improvement professionals to work together to create useful research summaries that include high-quality evidence from robustly designed impact evaluations, process evaluations and social validity research, which are distributed directly to head teachers. Second, school improvement professionals should work closely with schools to create and identify opportunities for staff to observe evidence-based programmes and practices in schools.

The present study revealed a greater depth of understanding of the factors that inform decisions when adopting school provision. Our findings suggest that when head teachers are choosing which programmes and practices to adopt, they will more often seek advice from and draw on the experiences of other head teachers using the programme within a school cluster and in other schools. However, there is promising evidence that some head teachers will seek out and/or follow recommendations from education professionals from external networks, for example advice regarding evidence-based practice from a school improvement adviser. This study has also confirmed that the compatibility of programmes with needs, values, culture and resources greatly influences adoption decisions. We suggest that these findings should be used by policy makers and researchers to improve transmission and adoption of more evidence-informed practices at the *knowledge* and *persuasion* phases of decision making in schools.

Chapter 5: Exploring the factors that affect adoption and implementation of the Good Behavior Game, following typical in-service training

Abstract

The Good Behavior Game (GBG) is a classroom management procedure with an extensive and promising evidence-base, is cost-effective, and requires minimal resources. However, like many promising programmes, it has failed to translate into every-day use in the classroom. To increase teachers’ skills and knowledge in evidence-based methods in-service training has been recommended. Other evidence suggest that training and coaching are effective methods for increasing the use of evidence-based provision. During a typical teacher in-service training day, we trained sixteen teachers from a cluster of schools in Wales, in the use of the GBG. For this study, we were interested to explore the extent that teacher would implement the GBG, following training, and the factors that influenced decisions to adopt or reject the procedure. Eleven teachers from nine primary schools, one secondary school, and one special school consented to take part in our study, which consisted of questionnaires, observations and interviews. We aligned our results with the concepts of Rogers’ (2003) diffusion of innovations theory, and found that even though knowledge was good, and attitudes towards the GBG were positive, only two teachers adopted the GBG. The remaining teachers either immediately rejected (*n* = 3) or trialled and rejected the programme (*n* = 4). The decisions of two teachers were unknown (*n* = 2), and none of the teachers in the study sustained use. Other findings from the study revealed that competing priorities and programmes facilitated decisions to reject the GBG. The findings from this study closely align with existing research and adoption of innovation theory, and advance the knowledge-base on how to increase the use of evidence-based programmes in schools. We suggest that more in-service training coaching will help to increase the uptake of evidence-based methods in schools.

For effective teaching and learning to take place, an orderly classroom is generally required. Frequent class disruptions from problematic behaviours have been found to result in reduced teaching and learning time (Jennings & Greenberg, 2009). A report published by the Office for Standards in Education (Ofsted; 2014) indicates that the most common form of problematic behaviours are low-level, high-frequency disruptions, which often result in decreased learning time for the whole class. For example, the report found that pupils are potentially losing up to an hour of learning every day because of low-level disruptions in classrooms, which is equivalent to 38 day of teaching (Ofsted, 2014). The most common behaviour challenges reported are talking unnecessarily, calling out, failing to listen to instructions, a lack of respect for each other and staff, and a lack of the right equipment (Ofsted, 2014; Reid, 2009). Findings from the National Behaviour and Attendance Review (NBAR) in Wales, has also found that continuous low-level disruption can have a serious and cumulative effect on pupils, and also on school staff (Reid, 2009). For example, behavioural disruptions in a classroom often interfere with pupils' ability to learn, and also with a teacher's ability to teach effectively, which can negatively impact the long-term academic outcomes of all learners in a classroom (Blank & Shavit, 2016; Ning et al., 2015; Organization for Economic Co-operation and Development [OECD], 2010), and lead to teacher stress, burnout, and lowered self-efficacy (Bradshaw et al., 2010; Browers & Tomic, 2000).

Efforts to teach pro-social behaviours in classrooms are now more commonly referred to as social and emotional learning (SEL) (Durlak et al., 2011). It is important that schools implement the most effective methods to develop pupils' social and emotional skills. Not only can it positively impact classroom behaviour, but it also fosters positive pupil and teacher relationships, enhances school connectedness, and improves educational outcomes (Durlak et al., 2011). The adoption of proven programmes also ensures that teachers' time and school budgets are used wisely (Embry, 2002; Gorard, See, & Siddiqui, 2020). In recent years, there has been a marked increase in the number of evidence-based programmes available that teach a range of social and emotional skills (Durlak et al., 2011; Goldberg et al., 2019). Through evidence repositories such as the Education Endowment Foundation (EEF) (https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects), teachers now have greater information about these programmes. An example of a cost effective and simple to use evidence-based classroom SEL procedure is the Good Behavior Game (GBG) (Durlak et al, 2011; Flowers et al., 2014).

The GBG is an interdependent group contingency procedure that was developed by Barrish, Saunders, and Wolf (1969) to reduce the disruptive behaviours (out-of-seat and talking-out behaviours) of students in elementary (primary) school settings. The GBG has been described as a behaviour modification technique that is both simple to use and cost-effective (Embry, 2002). The GBG has a strong theoretical base and extensive evidence supporting its use in a range of school contexts and with a variety of populations (see Bowman-Perrott et al., 2015; Coombes et al., 2016; Flower et al., 2014; Joselyn et al., 2019; Kellam et al., 2011; Nolan et al., 2014). The procedure allows teachers to engage in a number of behavioural management techniques such as the teaching of classroom expectations or rules (e.g., raise hand to speak); the acknowledgment of prosocial behaviours; and the reinforcement of behaviours through praise, reward, and feedback (Flowers et al., 2014). Much of the evidence supporting the use of the GBG has aimed to decrease disruptive behaviour in primary school classrooms, although it can also be used in secondary classrooms. Since its development, multiple group design and single-case design studies conducted nationally and internationally have shown the GBG to be effective in reducing challenging and disruptive behaviours and increasing pro-social behaviours in the classroom (e.g., Bowman-Perrott et al., 2015; Coombes et al., 2016; Flowers et al., 2014). Specifically, findings from two reviews report moderate to large effects on a variety of challenging behaviour (Bowman-Perrott et al., 2015; Flower et al., 2014). Flower et al. also report that reductions in disruptive behaviours from using the GBG result in increased teaching and learning time. The strategy has been included in a number of national and international registries and research repositories such as: the Early Intervention Foundation’s Guidebook (http://guidebook.eif.org.uk/), the Substance Abuse and Mental Health Services Administration: National Registry of Evidence-based Programs and Practices (https://www.samhsa.gov/resource-search/ebp); and has been certified as a ‘promising’ programme by the Blueprints for Healthy Youth Development database (http://blueprintsprograms.com/). However, recent evidence from a large-scale trial in the UK, commissioned by the EEF, reported equivocal results (Humphrey et al., 2018). Humphrey et al. found no evidence that the GBG behaviour was effective, but they did find tentative evidence that boys at-risk of developing conduct problems benefitted from the game. Furthermore, the trial was impacted by implementation issues, which are discussed later in this evidence review.

Like many evidence-based programmes and strategies, research suggests that the GBG has failed to translate into UK classrooms, and many teachers are unaware of its existence (Gorard, 2020; Stormont, Reinke & Herman, 2011). Gottfredson and Gottfredson (2002) report that many schools do not adopt evidence-based prevention programmes. Similarly, a report by Pegram et al. (2022 – Chapter 3) finds that much of school provision lacks a robust evidence-base. Even though research evidence and information about ‘what works’ is now more widely available, Nelson et al. (2017) and Walker et al. (2019) suggest that evidence repositories such as the EEF only have a small to moderate influence on the provision that schools use. Across the research literature there is a general consensus that teachers view much of the information as complex, often contradictory and irrelevant, (Hemsley-Brown & Sharp, 2003; van Schaik et al., 2018; Willingham & Daniel, 2021). Instead, teachers will seek advice and more practical and relevant information from staff in schools. (Chapter 4; Nelson et al., 2017; Walker et al., 2019). According to findings from Chapter 4 it was common for head teachers to visit a school to observe a programme in use. However, Chapter 4 also found that local education professionals, particularly the school improvement adviser (from the regional school improvement service in North Wales), were considered a trusted source for evidence-based recommendations.

It is also suggested that educators lack the time and the research literacy skills to access, comprehend and apply evidence effectively. There have been efforts from researchers to find better ways to mobilise research evidence for educator use. For example, school leaders should create a structure within schools that allows teachers time to engage in research and evidence (Brown & Zhang, 2016; Coldwell et al., 2017; Nelson & O'Beirne, 2014; Nutley et al., 2007), researchers should create more accessible summaries of research that clearly and concisely present the evidence that educators require (Gorard et al., 2020; Nelson & O'Beirne, 2014; Slavin, 2020), and, intermediaries should provide opportunities for teachers to learn about best-practice methods through in-service training and coaching (Nelson & O'Beirne, 2014).

The factors that more often inform educator decisions when choosing the programmes and practices to adopt closely align with the concepts of Rogers' (1995, 2003) diffusion of innovations theory. Diffusion of innovations theory defines how new ideas or practices (innovations) spread (diffuse) through a social system. According to Rogers, decisions to adopt new ideas and practices within an organisation are largely influenced by communications between members of a social network, and these members are referred to as *'change agents'*. Within the context of an education system, change agents are commonly staff members in schools (internal networks), but can also be members of external networks that are working towards a common goal. For example, a local school improvement advisor or educational psychologist would be considered change agents. According, to Rogers' (2003) theory, the innovation-decision process moves through five phases. *Knowledge* – an adopter becomes aware of an innovation and how it functions. *Persuasion* – an attitude is formed towards the innovation. *Decision* – adopter chooses whether to adopt or reject an innovation. *Implementation* – an innovation is put into use and may be modified by the user. C*onfirmation* – use of the innovation is sustained if decisions to adopt are reinforced.

Understanding the factors that facilitate and hinder the process of diffusion may help to increase the uptake of evidence-based programmes in real-world settings (Durlak & DuPre, 2008; Rogers, 2003). Rogers posits that during the pre-implementation phases of the model, adopters form attitudes towards an innovation that are strongly influenced by its perceived characteristics. Five characteristics have been proposed that predict adoption, which are: *relative advantage* (degree to which an innovation is perceived as better than the existing one), *complexity* (degree to which the innovation is perceived as difficult to use and understand), *compatibility* (degree to which an innovation is compatible with existing needs and values), *trial-ability* (degree to which the innovation can be tested), and *observability* (degree to which the innovation can be viewed in action). Durlak and DuPre (2008) conducted a review of the implementation literature and identified 23 factors that should be considered when attempting to implement innovations into real-world settings, some of which align with diffusion theory, including the compatibility, perceived need, and perceived benefits of an innovation, and the extent that an innovation can be adapted to fit. Durlak and DuPre also discovered other important factors such as the availability of funding, impact of leadership to drive the process, and training and further support/coaching to develop skills and promote self-efficacy. In a more recent review of the diffusion of innovation literature, relative advantage, complexity, compatibility and observability were identified as the most influential characteristics (Wisdom et al., 2014).

Unfortunately, even when programmes are adopted into schools, they are infrequently implemented as intended, and are modified by staff to better fit context (Durlak & DuPre, 2008). A consistent finding across the implementation literature suggests that programmes not implemented as intended often lead to unsuccessful outcomes (Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008; Sklad et al., 2017). A meta-analysis of universal SEL interventions found that programmes implemented with high levels of fidelity produced greater effects compared to those with low fidelity (Durlak et al., 2011). A recent evaluation in the UK, which reported null effects from the GBG on pupils' disruptive behaviour, found that the GBG was played less frequently and for shorter periods of time than recommended, implementation levels across teachers were moderate but varied, and 24 per cent of participants ceased implementation before the study concluded (Humphrey et al., 2018). Notably, Humphrey et al. found that the levels of implementation fidelity were not associated with pupil outcomes, but pupil responsiveness was. The same authors report that implementation issues arose from teachers' dissatisfaction with the time and effort involved, competing priorities, and problems with pedagogical fit (Humphrey et al., 2018). Other evidence across the GBG literature suggests that teachers who view the classroom management procedure as acceptable and suitable are more likely to implement the GBG with greater levels of fidelity (Wehby et al., 2012). Wehby et al. also found that greater levels of training and coaching were associated with greater fidelity. Dijkman, Harting and van der Wal (2014) found that the compatibility of the procedure with existing methods and needs facilitated decisions to adopt, and the presence of competing programmes influenced decisions to reject. Lastly, and in contrast to other GBG research (e.g., Flowers et al., 2014), Coombes et al. (2016) found that implementers felt burdened by the time required to play a game and with finding the appropriate opportunities to fit the game in.

Nevertheless, Durlak and DuPre (2008) postulate that highly structured and manualised programmes may necessitate greater levels of fidelity, whilst programmes without these features may be more conducive to contextual adaptions. The same authors report that implementation levels rarely achieve 80 per cent across all users in a study, therefore adaptions are inevitable (Durlak & Dupre, 2008). Advocates of the GBG report that there is flexibility in adherence to the GBG protocol, which means an implementer can modify the game to better fit the needs of a classroom, and still have positive effects. For example, class rules and rewards can be chosen, and length of play can be changed (Embry, 2002; Flower et al., 2014).

There is a growing evidence-base of promising SEL programmes, and this includes the GBG (Durlak et al., 2011). Based on the extensive evidence-base, the classroom management procedure has been described as cost-effective and relatively simple to use, and requires minimal resources to implement (Embry, 2002; Flowers et al., 2014). However, like many evidence-based methods, the GBG has failed to establish its use in UK schools, and many teachers are unaware of its existence (Gottfredson & Gottfredson, 2002) and, importantly, often use strategies that lack robust evidence (Pegram et al., 2022 – Chapter 3). To increase the use of evidence-based programmes in a cluster of schools in Wales, the local school improvement advisor disseminated information about the GBG to head teachers and invited them to select two members of staff to be trained in the use of the procedure during a typical in-service training day. With the exception of staff time, there was no impact on schools' resources, and they were under no obligation to adopt and use it. For this study we were interested to understand the extent to which teachers would implement the GBG following a typical in-service training, and the factors facilitating and hindering adoption-decisions. The findings from this study will also contribute to the knowledge-base on how to increase the use of evidence-based programmes in schools and the existing GBG literature. We asked the following research questions:

1. To what extent will teachers implement the Good Behavior Game following a typical in-service training day?
2. What factors influence decisions to adopt or reject the Good Behavior Game?

Method

Context

This study was conducted with the teachers from an established cluster of schools in Wales, which consisted of two secondary schools (English medium), seven primary schools (one Welsh medium & six English medium) and one special school (English medium). With the intention of moving towards a more evidence-informed practice, the cluster had previously entered into a research partnership with the Regional School Effectiveness and Improvement Service for North Wales (GwE) and the Collaborative Institute for Education Research, Evidence and Impact (CIEREI), Bangor University, to undertake close-to-practice research projects.

Good Behaviour Game training

The local school improvement adviser from GwE disseminated information about the GBG to head teachers and invited them to nominate up two teachers from any year group to attend a two-hour GBG training session. The head teachers were informed that all training and accompanying resources would be provided free of charge, and that schools were under no obligation to adopt the GBG. It was also agreed that at the end of the training session we would invite teachers to take part in a study in which we planned to evaluate adoption and implementation of the GBG.

Two hours was set aside during a staff in-service education and training day (INSET) in February 2018 to train teachers in the use of the GBG as part of a typical in-service training session. INSET is the term used to refer to a wide range of training activities that take place in school and contribute to the development of teachers already working in the profession (see Craft, 1996; Hargreaves & Fullan, 1992). Sixteen teachers from 10 schools (two English medium secondary schools, six English medium primary schools, one Welsh medium primary school and one English medium special school) attended the session, which took place in a classroom in one of the cluster schools. Training was delivered by the first author and included a step-by-step description of playing the GBG, modelling of the procedure, and the latest empirical evidence. Throughout training teachers asked questions and engaged in general discussions about how to play the game, types of rules to set, and rewards to provide. Post training, teachers were provided with copies of the training slides, and a GBG evidence summary produced by the Early Intervention Foundation (https://guidebook.eif.org.uk/programme/the-good-behaviour-game). Teachers were asked to indicate on a training evaluation form if they would like any further training or coaching in the classroom, and as a result additional coaching was provided to one teacher.

At the end of the GBG training session, we invited the teachers to take part in a study to evaluate implementation of the GBG and provided them with information sheets that detailed the study, and consent forms asking for signed consent for us to collect and analyse data.

Good Behaviour Game

We chose to train teachers in the procedure that awards points for following the rules rather than violating the rules, which departs from the original procedure used in the study by Barrish et al. (1969). This more positive variation of the GBG uses reward to reinforce desired classroom behaviour and has been found to positively impact disruptive and on-task behaviour (see Grove & Austin, 2017; Wright & McCurdy, 2012). The main features of the GBG procedure are outlined in Table 1. Teachers were advised to play the game at least three times a week.

Table 1. *An outline of the procedure for the Good Behavior Game*

	Action
Preparing for the game	1. Teach pupils 3-4 expectations (rules) for good behaviour and ensure they are simple, positively framed, and visible. 2. Divide pupils into at least two teams. 3. Plan how long you will play the game (approx. 20 minutes) and use a timer to keep track. 4. Decide how the game will be won (e.g., team with most points wins). 5. Decide what reward you will give to the winning team(s). 6. Remind pupils of the rules before playing the game and of the reward to be earned for following them.
Playing the game	1. Announce the start of the game and set the timer. 2. Award points to the teams who meet expectations and ensure points are visible. 3. Praise rule following. 4. At the end of the timing announce the winning team(s) and provide a reward(s).

Ethics

This research study was conducted in accordance with the ethical approval granted by the School of Educational Sciences Research Ethics Committee at Bangor University (ref: 18-19). All participants consented to participate and were informed of their rights to withdraw at any point. Before commencing the interviews, participants were provided with examples of the questions. To protect anonymity, identifiable information such as school and teacher names are not included.

Study design

We used a mixed methods design that consisted of questionnaires, semi-structured interviews and observations, and evaluated the findings using Rogers' (2003) theory as an organising framework. See Figure 1 for an adapted illustration of Rogers (2003) innovation-decision process, which includes the concepts that we considered.

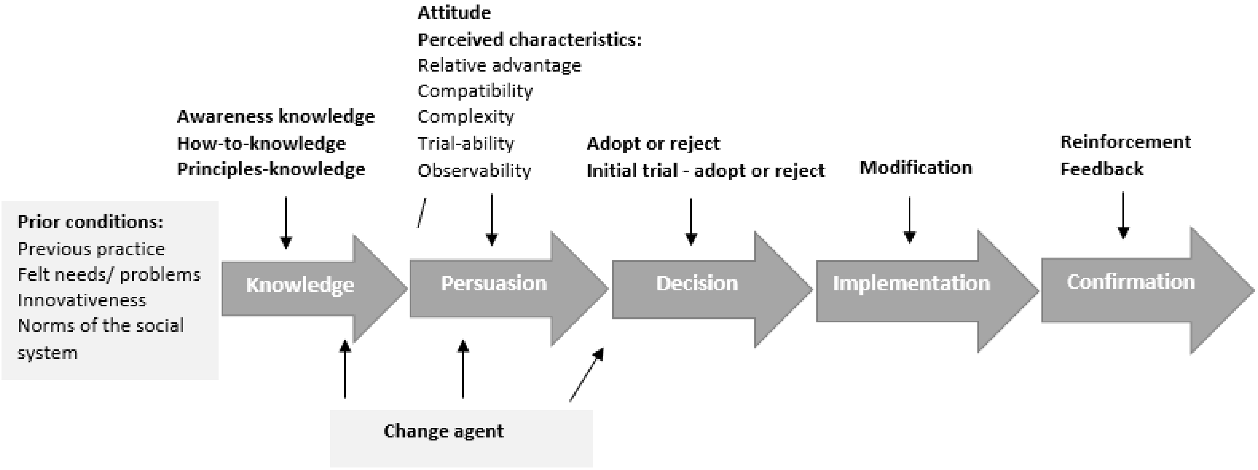

Attitude
Perceived characteristics:
Relative advantage
Compatibility
Complexity
Trial-ability
Observability
Awareness knowledge
How-to-knowledge
Principles-knowledge
Adopt or reject
Initial trial - adopt or reject
Modification
Reinforcement
Feedback
Prior conditions:
Previous practice
Felt needs/ problems
Innovativeness
Norms of the social system
Knowledge
Persuasion
Decision
Implementation
Confirmation
Change agent

Figure 1. Theoretical framework based on Rogers (2003) innovation-decision process

Participants

Twelve members of staff from eight of the schools consented to take part in the study. However, one teacher moved schools soon after the training session and we removed their data from the study. Table 2 shows the characteristics of the eleven teachers that participated, and Table 3 shows the participants' current classroom management strategies.

Table 2. *Characteristics of the teachers included in the study*

	Variable	*n*
School type	Primary	9
	Secondary	1
	Special needs	1
Staff role	Teacher	9
	HLTA*	2
Gender	Female	10
	Male	1
Years teaching	1 - 5	5
	6 - 10	2
	11 - 15	1
	16 and over	3

*HLTA = higher level teaching assistant

Table 3. *Displays the behaviour management techniques or strategies currently used by the teachers*

School type	Classroom management strategy(ies)
Primary	
1	Pivotal
2	Pivotal
3	Pivotal
4	Assertive discipline
5	Good to be green, dojos
6	Dojos, table points
7	Dojos, card system, raffle
8	Pivotal, positive rewards system, learn grow achieve
9	Assertive discipline, traffic lights, dojos
Secondary	
	Behaviour Watch
Special needs	
	Pivotal

Outcome measures

A demographic questionnaire was completed by 11 teachers. This was used to established characteristics such as role, gender, school type, years teaching and current method or strategy for managing disruptive classroom behaviour (see Table 2 and 3).

Eleven teachers completed a 13-item bespoke workshop training evaluation form, which was used to evaluate the training content and materials, teachers' knowledge towards the GBG, and if additional training was required. In particular, we were interested to understand if teachers had acquired adequate knowledge about the aims of the GBG *(principles-knowledge)*, how to use it *(how-to-knowledge)*, initial attitudes towards the GBG *(attitude)*, and if adequate knowledge had been acquired. The training evaluation form uses a 5–point Likert-type scale, which ranges from 1 (strongly disagree) to 5 (strongly agree).

We administered an adapted version of the Intervention Rating Profile ([IRP-15] Martens et al., 1985) to the teachers that played the GBG in the classroom. The remaining teachers that did not trial the GBG did not complete this measure because they had no experience in using the GBG to accurately rate it. The IRP-15 was used to evaluate teachers' *attitudes* towards the GBG. The IRP-15 uses a 6–point Likert-type scale, which ranges from 1 (strongly disagree) to 6 (strongly agree). A high score on the scale is associated with an acceptable intervention. The internal consistency of the instrument has been reported to be 0.89 (Tarnowski & Simonian, 1992). We adapted the items to reflect the application of the strategy to a group of pupils and to the present tense.

Teachers were also provided with a series of weekly logs to determine the frequency of GBG games played and the duration of each game (dosage). See Appendix B, C & D for examples of the measures.

Observations

To evaluate implementation fidelity at the *implementation phase* of the model, we conducted a structured observation with one teacher that was implementing the GBG. We created a data collection rubric to assess the extent that the GBG procedure had been adhered to (fidelity). The observation instrument was based on the PAX GBG rubric by Schaffer et al. (2006), and the GBG implementation rubric by Marchese et al. (2017). The rubric consisted of eight items that were completed with a mark to indicate "yes" or "no" if a step was observed or not. Example items included (1) pupils are placed into teams, (2) rules are displayed and/or pupil are reminded of rules, and (3) Rules are simple to understand and framed positively (see Appendix E for the rubric). A score of 1 was given to items marked "yes" and a score of 0 was given to items marked "no". A higher score on the rubric (total scores range from 0 to 8) is associated with high fidelity to the GBG procedure. During the observation, the teacher was asked to play the GBG for 5-10 minutes whilst the first author observed and completed the rubric. As only one observation was completed, reliability of the fidelity measure was not assessed.

Interviews

We conducted interviews with nine teachers and were unable to make contact with the remaining two teachers that worked in the same school. A semi-structured interview guide that included open-ended questions and prompts was prepared by the first author in alignment with the key constructs of Rogers' (2003) theory. The aim of the questions was to gain insight into information across the five phases of the model (*knowledge, persuasion, decision implementation and confirmation)*. We aimed to understand if participants were aware of the GBG prior to training *(awareness knowledge)*, reason for attending training *(prior conditions)*, attitudes towards the GBG, *(attitude and perceived characteristics)*, if the GBG had been adopted, put into use and sustained, and why *(decision, implementation and confirmation*); and, finally, if anyone influenced their decision (*persuasion and decision phases)*. To extract more detail, we used targeted prompts based on the constructs of the theoretical framework (e.g., *compatibility, relative advantages,*).

Each interview was conducted in school during school hours. Before commencing, the example interview questions were read to the participants. The interviews were audio recorded using a voice recorder and lasted approximately 15 minutes in length. Simple notes were made during each interview, and these were read to the participant at the end so they could add any additional comments and/or corrections.

Data collection procedure

We collected data across three main time points. At the end of February 2018, immediately after training, teachers completed a demographics questionnaire, a training evaluation questionnaire and were provided with a series of weekly GBG dosage logs (time-point 1). In early March 2018 we sent email invites to the teachers to determine if they had adopted the GBG and to organise a suitable day and time to conduct an interview. Throughout March and April 2018 interviews were conducted, the IRP-15 was completed, and a classroom observation was conducted with one teacher that was still using the GBG (time point 2). In July 2018, we planned to visit the teachers to collect dosage data and conduct final structured classroom observations with those still using the GBG (time-point 3). However, though email correspondence we were notified that at time-point 3 no teachers had continued to implement the GBG. See Figure 2 for the frequency of teachers that participated in each of the data collection methods.

Table 4. *Frequency of teachers that participated in the data collection methods*

Time point 1 (February 2018)	Time point 2 (March-April 2018)	Time point 3 (July 2018)
Training evaluation form: *n* = 11	Interviews *n* = 9 IRP-15: *n* = 6 Observations: *n* = 1	Observations: *n* = 0 Dosage data: *n* = 0

Analysis

We analysed data from the IRP-15 and the training evaluation form using descriptive statistics. Two of the respondents left questions 2 & 6 unanswered on the IRP-15 so those questions were removed from the analysis of all data. In addition, a total score was calculated for the GBG fidelity rubric.

The interview transcriptions and an open-ended question from the training evaluation form (question 14) were analysed using a deductive approach to content analysis, which Hsieh and Shannon (2005) and Newby (2010) refer to as ‘directed’ content analysis. This approach was chosen because of its ability to systematically search for the presence of, and code for, certain words, or concepts within a dataset, which are derived from a theoretical framework (Hsieh & Shannon, 2005).

The data analysis was carried out by the first author who was familiar with Rogers’ (2003) theory. According to the method outlined by Hsieh and Shannon (2005), first the transcripts were read, re-read, and each unit of data (e.g., words and sentences) that appeared to represent the applicable constructs were highlighted. The highlighted data were then coded using Rogers' (2003) theoretical constructs (e.g., *felt need/problems, relative advantage, and compatibility*). Lastly, a reanalysis of the data and codes was conducted. For any text that could not be categorised, we adopted an inductive analysis to ensure that new constructs and meaningful data were also coded.

The preliminary data categorisation was checked by an independent reviewer, who was made familiar about the framework and concept of Rogers' (2003) theory, and inter-rater reliability was reached on all of the data.

Results

The adoption-decisions for each member of staff (*n* = 11) and the total scores for teachers for questions 11 & 12 from the training evaluation form are displayed in Table 5. Scores indicate that all teachers had good levels of 'principles knowledge' and 'how to knowledge' of the GBG, following training. The mean total scores for each question on the training evaluation form are displayed in Table 6.

Table 5. *Adoption decisions and total scores for teachers for question 11 and 12 from the training evaluation form*

School type	Staff role	Adoption decision	Total score
Primary	Teacher	Adopt	8
Special needs	Teacher	Adopt	8
Primary	Teacher	Active reject	8
Primary	Teacher	Active reject	7
Primary	Teacher	Active reject	8
Primary	HLTA*	Active reject	10
Secondary	Teacher	Reject	10
Primary	HLTA*	Reject	9
Primary	Teacher	Reject	10
Primary	Teacher	Unknown	8
Primary	Teacher	Unknown	8

Note: Active reject = adopter trials an innovation and then reject it, total scores range from 2 – 10, *HLTA = higher level teaching assistant

Table 6. *Mean scores for the questions from the training evaluation questionnaire*

Questions from the training evaluation form	Mean scores
The training session was well organised	4.45
The content of the session was covered in the time available	4.64
Trainers provided all that I needed to complete training tasks	4.18
The presentation slides were relevant, clear and useful	4.18
The handouts were relevant, clear and useful	4.36
Trainer presentations were clear and concise	4.27
Trainers demonstrated practical skills and knowledge	4.09
Trainer feedback was clear and concise	4.18
I learned from the presentation	4.27
I learned from the group discussion	4.27
I know what ‘Good Behaviour Game’ aims to do	4.36
I know how to use ‘Good Behaviour Game’ in the classroom	4.18

Note: Scores range from 1 (strongly disagree) to 5 (strongly agree).

Table 7 displays the scores from the IRP-15 for the teachers that chose to adopt and actively reject the GBG. Higher scores suggest that the GBG is perceived as an acceptable classroom management strategy.

Table 7. *Total acceptability scores on the IRP-15 for the teachers that adopted and actively rejected the GBG*

School type	Staff role	Adoption decision	Total score
Primary	Teacher	Adopt	65
Special needs	Teacher	Adopt	76
Primary	Teacher	Active reject	65
Primary	Teacher	Active reject	60
Primary	Teacher	Active reject	63
Primary	HLTA*	Active reject	72

Note: total scores on the IRP-15 range from a possible 6 to 78, HLTA = higher level teaching assistant

One teacher continued to implement the GBG at time-point 2. The total implementation fidelity score from the one observation was 6 (75%), which suggests that the GBG was implemented with high fidelity (possible scores on the rubric range from 0 - 8). See Table 8 for the results from the observation rubric.

Table 8. *Results from the observation conducted with one teacher*

	Yes	No
1. Pupils are in teams	x	
2. Rules are displayed or pupil are reminded of rules	x	
3. Rules are simple to understand and framed positively	x	
4. Appropriate behaviours are monitored and tracked		x
5. Appropriate behaviour is verbally reinforced		x
6. The game is timed	x	
7. Winner/s announced	x	
8. Rewards are delivered	x	

The following section provides a summary of selected findings from our qualitative exploration with the members of staff. The data and chosen excerpts are discussed using Rogers' (2003) theory as an organising framework

Prior conditions

Felt needs/problems

Data show that four participants (1 adopter, 1 rejecter, and 2 active rejecters) attended the GBG training because they had an existing need or problem that they thought the GBG might address: *"I had a particularly disruptive child in the class, and I thought it would be a good to get an idea of how else I could maybe improve his behaviour"* (rejecter).

Innovativeness

Two participants (one adopter and one active rejecter) highlighted their interest in trying out the GBG because it was a novel idea that might help pupils: *"I just think it's a fresh idea. Something I haven't tried before. And I would like to see how it worked, if I did use it in the future with older children, if I get the opportunity to"* (active rejecter).

Continuing professional development

Although not explicitly one of the constructs in Rogers' (2003) theory. One participant claimed that they attended the GBG training to develop their skills and knowledge*: "So I thought, this is just another thing that's a good opportunity. It was put across as a behaviour management thing, and it's like, it's another arrow for the quiver as it were"* (active rejecter).

Knowledge phase

Eight participants (*n* = 2 adopters, *n* = 2 rejecters and *n* = 4 active rejecters) had no 'awareness-knowledge' of the GBG prior to training, and one reported some awareness: *"Yes, I'd heard about it from a colleague, who had heard about it, or read about it"* (Rejecter). Data from the training evaluation form indicated that one participant (adopter) required some additional training in the classroom: "*I would like 1-1 training"*.

Persuasion phase

Relative advantage

Interview responses revealed that four participants (two active rejecters and two rejecters) rejected the GBG because other competing classroom strategies were perceived to be more advantageous than the GBG: *"We've got particularly well established sort of system in school that we tried to divert from 12 months ago. Yeah. And we've come back to it, we always come back to it"* (rejecter). *"We have quite a good behaviour policy in place and consequences and reward system in place that were effective, anyway. So I was more sceptical about whether it would make a big difference, doing something different to what I already do"* (active rejecter). Two participants also talked about the GBG having similar characteristics to their existing system, which reduced their motivation to change: *“Well, I already kind of used it anyway but in a different way... So I kind of just carried on implementing that”* (rejecter). Furthermore, comments revealed that two of the participants, who were from the same school, and both actively rejected the GBG, perceived the strategy to be ineffective for some pupils: *“The good behaviour game worked for a certain percentage of the class that I find quite like games. But then there was one or two individuals who would be like, well, I'm not taking part”* (active rejecter). Conversely, two participants referred to continuing to use the GBG because it was somewhat effective on disruptive classroom behaviour: *“It was another option I could use with this class to try and keep class behaviour to an acceptable level. With this class, as soon as you turn your back they just rise. And this (referring to the GBG) would work some of the time”* (adopter).

Compatibility

One participant explained that disruptions in the classroom were minimal, which meant that the GBG was superfluous to need: *“If I was in a class with children that had difficulties, then I could see that I would be able to use it in a one off class situation that had a need for it. We haven't got this great need, really”* (rejecter). Two participants felt the GBG was too advanced for pupils in the classroom: *“The age of the children that I work with* (ages 4 to 5), *they didn't really get the understanding of what we're trying to achieve with the game”* (active rejecter). It also appeared that the time of year to conduct the training (February) was not appropriate for the immediate adoption and implementation of new classroom management strategies. Four participants (two active rejecters, one adopter and one rejecter) suggested that the GBG would be more effectively implemented if adopted at the beginning of term in September: *“The primary issue was we didn't implement it in September, if we implemented it in September, the children would have been used to it and it* *would be part of the system"* (active rejecter). *"We didn't implement it only because we were halfway through year seven"* (rejecter).

Observability

The lack of opportunity to observe the GBG being used in a classroom was mentioned by two participants (one adopter and one active rejecter): *"I would have liked is to see it in action in the classroom, which I haven't unfortunately had the opportunity to do… It'd just be nice to see it in practice"* (active rejecter).

Change agent

One participant referred to discussing the GBG with another staff member who was also part of the study, and it seems that this may have influenced the adoption decision: *"I think, after discussing with (another teacher in the study) there's certain aspects of it within our school as a total as an entirety that wouldn't work because of the children that we've got"* (active rejecter).

Decision phase

Four participants indicated that they trialled the GBG with pupils in the classroom and subsequently decided to reject it: *"I did trial it, yeah, for a few days. I gave it as best I could"* (active rejecter).

Implementation and confirmation phases

Two participants chose to adopt the GBG and put it into use. However, at time-point 2, one participant had discontinued use: *"I did for it for a period of time. And then I think, rightly or wrongly, Estyn came along, and I felt it wasn't imbedded enough"* (adopter).

Both participants that adopted and implemented the GBG did not document the number of games played and duration, using the GBG rubric. Therefore, the frequency of delivery is unclear. However, interview data revealed that the staff member who continued to implement the GBG at time-point 2 reported that the game was not played regularly: *"I have done it again. We had a gap, then, we implemented it again. It's all on the board now, but it's a bit hit and miss when I use it… We're trying to stretch it a bit more. Over a longer period of time, rather than weekly"* (adopter).

Discussion

The GBG is a promising strategy that has a wealth of evidence supporting its effective use in the classroom (Flowers et al., 2014). However, as with many other evidence-based programmes, research suggests that the GBG has failed to translate into every-day use in the classroom (Gottfredson & Gottfredson, 2002). With the aim of increasing the use of evidence-based programmes in a cluster of schools, and to contribute to the knowledge-base on how best to increase the use of evidence-based methods in schools, we trained teachers in the use of the GBG during a typical in-service training day. For this study, we were interested to explore the extent that teachers would implement the GBG, following typical in-service training, and to understand the factors and perceived characteristics of the GBG that facilitate adoption-decisions. Our main findings suggest that all of the teachers' reported having good knowledge of the GBG, and the GBG was viewed as an acceptable procedure by the six teachers that completed the IRP-15. However, only two teachers chose to adopt the GBG, three teachers rejected the GBG immediately after training, four teachers trialled the GBG and then rejected it (active rejection), and for two teachers their decision was unknown. One teacher discontinued use of the GBG at time-point 2 (March and April), and by time-point 3 (July) none of the teachers continued to implement the GBG. We conducted an observation with the one adopting teacher that continued to use the GBG at time-point two and found that the GBG was implemented with high levels of fidelity (75%). However, interview data revealed that the same teacher played the game infrequently, which is a common finding across the implementation literature (see Durlak & DuPre, 2008; Humphrey et al., 2018).

In accordance with the concepts of Rogers' (2003) theoretical framework, our interview data revealed that '*felt needs/problems*' were one of the main reasons that both GBG adopters and rejecters attended training; '*relative advantage'*, '*compatibility'* and '*observability'* were the perceived characteristics of the innovation that facilitated decisions to reject; and the perceived '*relative advantage'* of the GBG facilitated decisions to adopt.

Research indicates that active forms of learning in a supportive environment such as modelling, role play and feedback are more likely to promote the acquisition of skills (Bell & Kozlowski, 2008), and our training included many of these elements. The evidence from the training evaluation form suggests that teachers acquired good knowledge of the aims of the GBG and how to use it. However, our findings revealed that typical in-service training in the use of the GBG did not lead to the adoption of the procedure by all teachers. For training to be effective, research suggests that re-training and/or coaching is required to continue developing skills and self-efficacy (Durlak & DuPre, 2008; Greenwood et al., 2003; Owen et al., 2021). According to findings from a study by Joyce and Showers (2002), only five per cent of teachers that attended training put the newly learnt strategies into practice. However, when training was followed by re-training and coaching 95 per cent of teachers implemented strategies. Similarly, Wehby et al. (2012) reported that greater levels of training and coaching in the GBG led to higher levels of implementation. Therefore, our findings suggest that typical in-service training, without re-training and additional coaching, is not an effective way of increasing the use of an evidence-based strategy in schools.

Contrary to the existing GBG literature (e.g., Domitrovich et al., 2015; Rogers, 2003; Wehby et al., 2012), our data revealed that high teacher-ratings of acceptability and suitability of the GBG did not lead to high levels of implementation for the majority of the teachers in this study that completed the IRP-15 (two adopters and four active rejecters). One adopting teacher rated the GBG as acceptable and suitable, and implemented the procedure with high levels of fidelity (75%). We did not collect observation data from the other adopting teacher as they had discontinued use of the GBG by time-point 2 of data collection, and the remaining teachers that rated the GBG as highly acceptable rejected it after the initial trail. Furthermore, none of the teachers in our study sustained use of the GBG. Upon reflection, it would appear that teacher-report ratings of social validity through the IRP-15 reflects a more collective judgement about the nature of the GBG rather than the teacher's personal perspective. As a result, high levels of social validity on the IRP-15 suggests that the two members of staff that adopted and four that actively rejected the GBG viewed it as an acceptable and suitable procedure for managing classroom behaviour, but did not necessarily perceive it as acceptable and suitable in the context of their own classroom with other competing demands. Furthermore, and notably, both adopting teachers in our study neglected to complete the implementation dosage checklists so we cannot report dosage data. However, interview responses revealed similar results to Humphrey et al. (2018) that teachers did not play the GBG regularly. Pegram et al. (2022 – Chapter 3) report that schools have a large number of programmes and strategies in use, particularly programmes that promote social and emotional skills. We suggest that our findings may also reflect a lack of capacity to adopt and use yet another procedure.

The data from our qualitative interviews support previous findings from Dijkman et al. (2014) and Humphrey et al. (2018) and indicate that most of the teachers who rejected the GBG viewed other competing classroom management strategies as more advantageous than the GBG, or too similar to warrant change (*relative advantage).* In line with Humphreys et al.'s findings, many of the teachers who rejected the procedure referred to its incompatibility (*compatibility*) with pupils in the classroom. Two teachers felt it was too advanced for Year 1 pupils and one felt that is was superfluous to need. Another pertinent compatibility issue for teachers was the time of year in which the training was conducted. Teachers felt that the

GBG should be introduced at the beginning of a school year in September. Finally, the data show that existing pupil need (*felt needs/problems*) was one of the main reasons that both GBG adopters and rejecters attended training, which was one of the factors associated with adoption-decisions in Pegram et al. (under review – chapter 4).

As outlined earlier in the discussion, we found that typical in-service training is not a good method for increasing the uptake of evidence-based methods, and we suggest it is likely that the lack of adequate follow-up training and coaching led to insufficient mastery of the skills required to deliver the GBG, as well as a lack of teacher motivation and self-efficacy (Dusenbury & Weissberg, 2017; Durlak & DuPre, 2008; Ransford et al., 2009). We also acknowledge the lack of opportunity to observe the GBG in a classroom setting, which according to Pegram et al. (under review – chapter 4) and Rogers (2003) facilitates adoption-decisions, and was referred to by two teachers during the interviews. We suggest that the lack of comprehensive training in the use of the GBG, and the lack of opportunity to observe it being used in a classroom-setting, may have influenced teachers' perceptions of the characteristics of the GBG. For example, we propose that without mastery levels of knowledge, teachers may have been unable to adequately appraise the relative advantage and compatibility of the procedure. We propose that the lack of re-training and coaching were the main factors that underpinned the teachers' decisions to reject the GBG, as well as a lack of capacity to adopt another innovation into use.

Limitations

The number of teachers that took part in this study is low and the results would not be expected to generalise to a wider population. However, our results closely align with the results from existing implementation research (e.g., Durlak & DuPre, 2008; Rogers, 2003), which increases the validity of our findings, and the possibility that these findings might generalise to similar education settings.

We believe an important limitation of this project was the time of year in which training took place. As discussed earlier, the data from the interviews suggests that teachers felt that training in the use of the GBG should have been conducted at the beginning of the academic year in September, to allow for the successful integration of the procedure. Lastly, we acknowledge the failure to collect long-term data to evaluate the potential up-take of the programme in the following academic year. According to the literature, it can take multiple years for an innovation to embed (Fixen et al., 2009; Rogers, 2003; Rohrbach et al., 2006).

Implications

To increase the uptake of evidence-based methods in schools (Durlak & DuPre, 2008; Greenwood et al., 2003), we recommend that schools receive more comprehensive training in evidence-based methods that includes opportunities to re-train and receive coaching. Owen et al. (2021) reports that even small amounts of coaching can have positive effects. According to Pegram et al. (2022 – Chapter 3), schools are using a large number of programmes and strategies, which impact teacher time and availability for training, as well as school finances. We suggest that schools consider refining provision to ensure that school budgets and time are spent more efficiently; and, to allow for more time for teachers to engage in comprehensive training to achieve mastery levels of knowledge and skills, which are also required to make informed adoption-decisions. Furthermore, we also propose that school improvement professionals work more closely with schools to create and identify opportunities for staff to observe evidence-based programmes and practices in use in schools.

For this study we trained teachers in the use of the GBG and found that only two teachers adopted the procedure, seven teachers rejected the programme, and two decisions were unknown. By the end of the study period, none of the teachers sustained use of GBG. Our findings suggest that typical in-service training alone is not enough to increase the uptake of evidence-base methods in schools. We also found that competing strategies and, importantly, the lack of re-training and coaching and opportunity to observe the GBG in use, facilitated decisions to reject the GBG.

Chapter 6: A cultural and contextual adaption of the PATHS programme to enhance implementation and acceptability

Abstract

Adoption and implementation of programmes in schools is complex, and often, evidence-based programmes perceived as incompatible are reject or implement with low fidelity. It is unlikely that programmes are implemented completely all of the time, thus, modifications are inevitable. Dependant on the extent of modification, changes to a programme and less than optimal deliver can result in diminished effects. The Promoting Alternative Thinking Strategies (PATHS) programme has a wealth of supporting evidence. However, some teachers perceive PATHS to be incompatible with UK schools. For this study, we worked with teachers from a cluster of schools in Wales to culturally and contextually adapt the PATHS programme. We were interested to explore, if the adapted programme would lead to impact on pupils' outcomes, positive attitudes, high levels of implementation, and intentions to sustain use. We also aimed to advance the evidence base by describing the process of adapting PATHS. Our participants were six teachers and 148 pupils from five primary schools and one special school. We evaluated the adapted programme using teacher-report measures of pupils' social and emotional skills, self-report measures of implementation, observations, and focus group interviews. We found that the adaption objectives were achieved, levels of implementation were high, teacher attitudes were positive, and evidence of small to medium causal-effects were observed. We also discovered that competing priorities had a negative impact on dosage, and competing programmes may prevent sustained use. Now more robust research is required to evaluate the efficacy of the programme. The process of adaptation and results from this study highlight the need for more rigorous process evaluations prior to commencing efficacy and effectiveness trials.

There is an expansive evidence base that supports the role of social and emotional competence in healthy human development. Social and emotional competence can be described as the ability to recognise and manage emotions, resolve problems effectively, show empathy, and establish and maintain positive relationships. Through social and emotional learning (SEL), social and emotional competencies are developed (Domitrovich et al., 2017). In addition to developing social and emotional skills, evidence suggests that SEL has positive effects on wellbeing, relationships, academic performance, is essential for success in work and life, reduces the risk of emotion dysregulation, distress, and maladaptive behaviours (Durlak et al. 2011; Greenberg, Domitrovich, & Bumbarger, 2001). According to the literature, the development of social and emotional skills through SEL is equally important for children and young people as the development of academic and practical skills, taught in an educational setting (Organisation for Economic Co-operation and Development, 2017; Taylor et al., 2017).

There is a common understanding that social and emotional learning can, and should, be taught universally by teachers in schools (CASEL, 2003; Elias et al., 1997; Greenberg et al., 2003). In Wales, social and emotional wellbeing is one of the core areas of learning and experience, which is embedded throughout the new Curriculum for Wales (Welsh Government, 2017). In recent years, there has been a rise in the number of school-based programmes and strategies that teach social and emotional skills (Payton et al., 2008; Weare, 2010; Weare & Gray, 2003). Some programmes have a robust evidence base that demonstrates positive impact on pupils' skills, whilst others have preliminary evidence, or are simply untested (https://educationendowmentfoundation.org.uk/projects-and-evaluation/projects) (Gorard, See & Siddiqui, 2020). There have been three rigorous meta-analyses that have demonstrated the meaningful improvements that efficacious school based SEL programmes can have on pupils' social and emotional skills and mental health (Durlak et al., 2011; Sklad et al., 2012; Wigelsworth et al., 2016). Much of the evidence recommends that SEL is delivered frequently through classroom-based programmes and is embedded throughout the curriculum and school (Domitovich & Greenberg, 2000; Durlak et al., 2011). However, historically, personal and social education (PSE) in schools has been under-resourced, and under-valued compared to the teaching of academic subjects (Stallard et al., 2012).

Developed in the USA by Greenberg and Kusche (1993), PATHS is a school based SEL programme that teaches emotion understanding, emotion regulation and problem-solving skills. The programme is designed to be delivered by teachers and includes a series of lessons on topics such as identifying and labelling feelings, includes activities to promote the generalisation of skills, and parent materials to extend learning to the home. PATHS is supported by an extensive international literature base, which includes multiple randomised controlled trials (RCTs) (e.g., Domitrovich, Cortes, & Greenberg, 2007; Greenberg et al., 1995). More commonly, evaluations of PATHS have been conducted in the general education setting whereas research has been conducted with special education populations and found that the positive effects can be replicated (Kam et al., 2004; Greenberg et al., 1995). In the USA, PATHS has been identified as a ‘model program’ by both the Center for Study and Prevention of Violence (CSPV, 2006) and the Substance Abuse and Mental Health Services Administration (SAMHSA, 2011).

In 2011, an influential UK government report recommended the widespread adoption of PATHS based on its efficacy and multiple site applications (Allen, 2011). Nevertheless, the findings from research conducted in the UK has had mixed outcomes. For example, evidence from a small-scale quasi-experimental study reported significant and positive effects on a range of social and emotional competencies of pre-school pupils (Hughes & Cline, 2014), and a moderate-scale quasi-experimental studies reported significant improvements on pupils' mental health (Curtis & Norgate, 2007). A large-scale longitudinal RCT conducted in Belfast, Northern Ireland found weak but positive effects of PATHS on social and emotional competencies (Ross et al., 2011). However, a more recent large-scale cluster-RCT conducted in Birmingham, England found no statistically significant differences between participants in the PATHS group compared to the control on measures of mental health (Berry et al., 2015). The most recent evidence from a large-scale cluster-RCT in the UK conducted by Humphrey et al. (2018) indicated that PATHS had small but significant improvements on pupils' social skills and psychological wellbeing, but no impact on academic outcomes. Furthermore, the authors report variable implementation, which they suggest may have impacted the results. An economic analysis of the data collected further suggested that PATHS has the potential to be cost effective when used to promote social and emotional skills (Turner et al., 2020).

The rationale for adopting and using evidence-based provision (i.e., programmes that have already been found to be effective through robust research) in schools is that pupils are exposed to programmes and strategies that have evidence of positive impact on learner outcomes, and staff time and school budgets are used more efficiently (Gorard, 2020). However, evidence indicates that many schools do not use evidence-based programmes (Gottfredson & Gottfredson, 2002; Pegram et al, 2022 – Chapter 3). According to diffusion of innovations theory (Rogers, 2003) and findings from interviews with head teachers, programmes deemed contextually compatible are more likely to be adopted and implemented in schools (Pegram et al., under review – Chapter 4). Understandably, schools will adopt new ideas and programmes if they perceive them to be compatible with the needs of pupils, and the needs of the school. For example, head teachers will adopt programmes that they perceive to be advantageous, easy to use, and cost and time efficient. Crucially, school staff may reject evidence-based programmes if deemed a poor fit, and, if existing programmes and methods are perceived to be superior (Bumbarger, & Perkins, 2008; Chapter 4; Rogers, 2003).

However, a study conducted by Pegram et al. (2022) indicates that despite some conditions that enable the uptake of evidence-based strategies in schools, much of existing school provision has weak causal evidence or is untested.

The adoption and implementation processes are complex, and often when evidence-based programmes are adopted they are not implemented with fidelity and/or are modified by teaching staff to fit the context. It is common for school staff to use components of programmes that best fit with teaching practice and need (Durlak, 2016; Greenberg et al., 2005; Rogers, 2003). Issues such as complexity, limited resources (time & money), diverse pupil needs, existing culture and beliefs, and insufficient knowledge and training are some of the factors that prompt modifications to programmes (Domitrovich et al., 2008; Durlak, 2016; Rogers, 2003). Although often not assessed, the evidence suggests that the implementation of SEL programmes in schools is inconsistent, and this has been found to compromise effects (Durlak & DuPre, 2008; Durlak et al., 2011). SEL programmes implemented with low *quality* (the manner and level of skill in which a programme is delivered), *fidelity* (the degree to which the essential key components of a programme are maintained), and *participant responsiveness* (the extent that participants are engaged and respond to a programme) are less likely to have a positive impact on pupils' social and emotional outcomes. However, optimal *dosage* (frequency of delivery) does not necessarily have to be achieved to produce the desired effects (Domitovich & Greenberg, 2000; Durlak & DuPre, 2008). Nevertheless, the findings across studies that have explored the impact of PATHS implementation on social and emotional outcomes are somewhat conflicting (e.g., Berry et al., 2015; Faria et al., 2013; Humphrey, Barlow & Lendrum, 2018; Humphrey et al., 2016). Most recently, Humphrey et al. (2018) found through observations in UK schools that high levels of implementation quality and responsiveness were associated with greater social and emotional skills; and, contrary to the common narrative dosage was associated with lower social and emotional skills.

Adaptation is commonly defined in the implementation science as the intentional modification of a programme to achieve a better fit with a new context. Adaptations of a programme might include changes to the content of the intervention and/or the way it is delivered, or may include cultural adaptions for different populations. According to a recent review of the adaption literature for health interventions, the rationale for adaption is that it saves time and costs; addresses cultural and contextual differences; enhances cultural and contextual relevance; and ensures attractiveness, reach, and salience (Movsisyan et al., 2021). Arguably, the process of adaption is in conflict with implementation fidelity (delivery as intended), and the mixed causal evidence may suggest that the extent of programme adaption should be carefully considered. As described by Embry & Biglan (2008), it is essential that the key kernels (i.e., the active ingredients in existing programmes) be retained, since it is the kernels that contribute to a programme’s effectiveness and efficiency. It is also suggested that adaptation and fidelity co-exist, and that adaptation is a component of implementation just like fidelity (Barrera et al., 2017). It is highly unlikely that a programme is implemented with 100 per cent fidelity all of the time, which means adaptations are inevitable. However, evaluations of implementation often focus on fidelity and dosage, sometimes on quality, and adaptations are infrequently explored (Durlak & DePre, 2008).

To date, the PATHS programme has been adapted in many diverse English and non-English speaking cultures nationally and internationally. Indeed, the PATHS materials have been culturally adapted by children’s charity Barnardo’s, and a trial of this was conducted in UK schools by Humphrey et al. (2018). Adaptations were made to the vocabulary, names and cultural references, and the structure and delivery of the original PATHS programme was retained. However, the process of adapting PATHS was not explained. Based on the evidence by Humphrey et al. (2018) and Humphrey et al. (2016), teachers that implement PATHS UK do so with high fidelity, quality, participant responsiveness and reach. However, the frequency of delivery (dosage) was reported at around half of the recommended dosage, which is a common finding throughout the PATHS implementation literature (Berry et al., 2015; Faria et al., 2013). Furthermore, insights from teacher interviews, conducted by Humphrey et al. (2018), suggest that PATHS UK was deemed compatible with the values and needs of some intervention teachers. However, several issues were also raised, suggesting that the cultural adaptation for PATHS UK may have been somewhat inadequate. For example, for many teachers the materials were appropriate, but others perceived them as insufficiently anglicised. Concerns were expressed about the length of lessons, and difficulties were reported with finding adequate time and space in the curriculum to deliver PATHS optimally. This led Humphrey et al., (2016) to suggest that low dosage rates may have contributed to the failure of PATHS to improve outcomes. Humphrey et al. (2016, p.23) also found that teachers made predominately ‘surface level’ adaptions to the programme. For example, changes to language and accommodations for pupil need. However, other more significant adaptations, such as the omission of lessons and concepts, were also discovered.

The process of adapting PATHS for cultural and/or contextual relevance is rarely systematically described in the literature (e.g., Hughes & Cline, 2015; Kam et al, 2011; Humphrey et al., 2018). This study identified only two other studies that described the process of adapting PATHS, and only one conducted a pilot evaluation of the newly adapted programme (Inam, Tariq & Zaman, 2015). Inam et al. (2015) described a cultural adaption of PATHS for use with Pakistani children, and described the process using the heuristic framework of adaptation (Barrera & Castro, 2006), and more recently Ferrer-Wreder et al. (2012) described a cultural adaption of PATHS for the Swedish context, using the Planned Intervention Adaptation (PIA) protocol. Typically, models of adaption include a series of steps, which include, but are not confined to, programme selection, the creation of a team of relevant stakeholders, activities to gather information, designing of the adapted programme, pilot research, further adaptions and refinements are made (Barrera & Castro, 2006; Ferrer-Wreder et al., 2012).

The following study was conducted with an established cluster of schools in Wales. The cluster included two English medium secondary schools, seven primary schools (six English medium and one Welsh medium) and one English special school. With the intention of moving towards a more evidence-informed practice, the cluster had entered into a research partnership with the Regional School Effectiveness and Improvement Service for North Wales (GwE) and the Collaborative Institute for Education Research, Evidence and Impact (CIEREI), Bangor University, to commission this and other close-to-practice research studies. Throughout the project, the first author who was embedded within the cluster worked collaboratively with the head teachers and senior leaders to develop and plan research.

We discovered through anecdotal discussions with head teachers that the cluster primary schools had previously adopted and used the PATHS programme based on recommendations from a local educational psychologist and the Child and Adolescent Mental Health Services (CAMHS). A report by Appleton (2000) also alludes to the uptake of PATHS in other North Wales schools. However, the teachers in the cluster had largely discontinued use of this evidence-based programme because they perceived it to be contextually incompatible. Similar to the views reported by Humphrey et al, (2018), staff perceived PATHS to be insufficiently anglicised, and somewhat complex and lengthy for a curriculum already at capacity. On account that PATHS has demonstrated positive impact on pupil outcomes, we agreed with senior leaders that we would work with teachers to culturally and contextually adapt the PATHS programme.

For this study, we aimed to understand if a culturally and contextually compatible PATHS programme would lead to more positive attitudes, higher levels of implementation, and intentions to sustain use of the programme. In doing so, we hoped to increase the use of an evidence-based programme, and contribute to the literature on the best ways to increase adoption and sustained implementation of evidence-based provision in schools more generally. Furthermore, we also aimed to advance the evidence-base by describing the process of adapting PATHS. Specifically, we were interested in answering the following five research questions:

1. What is the process of adapting PATHS?
2. Were the objectives of adaption achieved?

   The adaption objectives were to: a) sufficiently anglicise the programme content and materials and include opportunities to teach Welsh; b) reduce the requirement for twice-weekly dosage; and c) condense the content whilst retaining the key kernels (core ingredients) of the programme.
3. To what extent will teachers implement the adapted PATHS programme?
4. What are teachers' attitudes towards the programme?
5. What impact does the adapted programme have on social and emotional outcomes of Year 1 pupils?

Method

Ethics

This research study was conducted in accordance with the ethical approval granted by the School of Educational Sciences Research Ethics Committee at Bangor University (ref: 19-23). We obtained consent from the head teachers from seven schools to implement the adapted PATHS programme with all pupils in one Year 1 classroom. Consent was collected from the participating teachers, allowing us to use their data, and parental consent to use pupil data was obtained from the parents of the pupils participating in the study. All head teachers, teachers and parents were informed of their rights to withdraw at any point. To protect anonymity, identifiable information such as school, teacher and pupil names are not included.

PATHS® (UK)

The PATHS® programme is based on the Affective-Behavioral-Cognitive-Developmental model of development (ABCD). To promote social and emotional competence the model focuses on the developmental integration of affect, emotion language, behaviour and cognitive understanding (Greenberg & Kusche, 1993). The core components of the programme include a taught curriculum and generalisation activities and techniques, which are delivered to all children in a classroom. Curriculum packs contain lessons and send-home activities that cover topics such as identifying and labelling feelings, controlling impulses, reducing stress and understanding other people's perspectives; and supporting resources such as posters, books and puppets.

The PATHS® concepts are typically taught through direct instruction, discussion, and modelling using stories, and the lessons follow a common format that includes an introduction from the teacher (in which the lesson topic and objectives are introduced), a main activity (often built around a group activity or story), and a brief plenary/closure (in which learning is reviewed). Frequent prompts to elicit pupil responses and clarify learning are included throughout. The programme utilises a “spiral” curriculum model, in which topics and concepts are revisited. The units and lessons are developmentally sequenced, and new learning is linked to previous learning with the aim of increasing learners’ competence with each successive visit to a topic or concept.

Delivery of the PATHS® lessons to all children is generally undertaken as part of the normal class timetable and generalisation activities and strategies are implemented routinely throughout the school day. It is recommended that PATHS® lessons last approximately 30–40 minutes and is delivered twice-weekly throughout the school year. On average a curriculum packs contain 40 lessons (Kusché, & Greenberg, 2005).

Process of adapting PATHS

To adapt the PATHS® programme (PATHS Reception/Year 1), we chose to follow the heuristic framework described by Barrera and Castro (2006). We chose this framework because it covers the adaption of programme content and delivery, and has been previously used by Inam, Tariq and Zaman (2015) to culturally adapt PATHS. First, we created a team that consisted of the first author and a member of staff from each of the schools in the cluster who was either a well-being co-ordinator or had an interest in well-being. In total we recruited nine teachers to the group from nine of the cluster schools (six primary, one special needs, and two secondary schools. We then held four face-to-face meetings in July, September, and November 2018, and January 2019, in one of the cluster schools. The meetings included detailed discussions, led by the researcher, to identify the core programme components to retain, and the content and processes that required adaption, as well as drafting the programme content and the supporting materials. We held a final meeting in July 2019, after completion of the piloting of the adapted programme, to discuss the findings, further adaptions, and plans for adapting PATHS® for other year groups. Table 1 displays the process and actions taken to adapt the programme.

Table 1. *An outline of the systematic process of adaption*

Adaption stages (Barrera & Castro, 2006)	**Actions**
Create an adaption team	1. Developed a team of relevant stakeholders with an interest in adapting the PATHS programme
Information gathering	1. The group discussed the context, needs of pupils and teachers, and the current SEL approaches used 2. The researcher reviewed the literature for other evidence-based kernels for teaching SEL and dosage recommendations

	3. The group read through the PATHS curriculum to identify the core programme components to retain 4. The group discussed and evaluated the language, structure, materials, and frequency of delivery 5. The group discussed other activities that might be incorporated
Preliminary adaptation design	1. The researcher compiled the information and presented the ideas to the group 2. The group designed and drafted the lesson plans 3. The group created draft materials (e.g., feelings faces) and sourced additional resources (e.g., a list of available books) to support the lesson plans 4. The group reviewed the programme content and materials and made final changes 5. The group chose 12 lessons for a preliminary feasibility study 6. The researcher created the final programme resources for the feasibility study
Preliminary adaption test	1. The researcher conducted a preliminary feasibility evaluation of the adapted PATHS programme (described below)
Adaption refinement	1. Present data from the evaluation to the group 2. Discuss further adaptations and research

Preliminary feasibility evaluation

Intervention

For the pilot study, we chose 12 lessons from the adapted PATHS programme that were based on the following PATHS® topics: ‘knowledge and awareness of feelings and the feeling of others’ (which includes teaching the feelings happy, sad, angry and scared), ‘self-control’ and ‘prosocial behaviours’ (which includes the teaching of compliments, rule following, and behaviours associated with being a good friend). We retained much of the core content and format of the PATHS® programme and, rewrote the information into more succinct lesson plans and accompanying scripts; adapted the activities, language and cultural references; incorporated opportunities to teach the Welsh language; developed discretionary extension activities; and created supporting materials to replace all of the PATHS® materials. We chose not to include the PATHS ‘kid of the week’ component because the group felt this did not fit with beliefs, values and existing reward systems. Comparable to the original PATHS® programme materials, we included generalisation techniques and activities to be implemented throughout the school day. In contrast to PATHS®, we did not include materials for use at home with parents.

The adapted PATHS programme currently includes 12 lesson plans and supporting materials. The lesson plans outline the lesson objectives and success criteria, and include three required lesson activities, a list of discretionary extension activities, and a list of generalisation activities based on the objectives of the lesson (see Appendix F for examples of the lesson plans). The supporting materials include scripts to accompany the lessons, which are based on a direct instruction model-lead-test approach (Engelmann & Carnine, 1982), and used to systematically and explicitly teach the key concepts (an example can be found in Appendix G). Other supporting materials include pictures depicting feeling faces and pro-social behaviours (e.g., rule following and sharing), short stories, and a list of accessible books (see Appendix H for examples). To replace the PATHS® turtle technique, and teach self-awareness and self-management/control skills, we developed a script based on the ‘noticer’ skill from the DNA-V model by Ciarrochi and Hayes (2016). Based on evidence from PATHS research (e.g., Berry et al., 2016) we recommended that PATHS lessons be delivered once a week for approximately 30-40 minutes.

Trial design and participants

A multi-method approach was adopted for the preliminary feasibility trial of the adapted PATHS programme. We used structured observations and implementation checklists to assess components of implementation; focus groups to address suitability, acceptability and further adaptions; and a pre- and post-test quasi-experimental design to measure impact on pupils’ social and emotional outcomes.

We recruited seven schools (six primary schools and one special school) from the cluster to participate in the study. However, one of the primary schools withdraw from the study before pre-test data was collected. The participants were the teachers and pupils from Year 1 classes (aged 5 to 7 years). Five of the schools (four primary schools and one special school) had two Year 1 classes, and one primary school had one Year 1 class. The head teachers from each of the schools selected one Year 1 classroom to be assigned to an intervention group (adapted PATHS programme) and the other Year 1 classroom to a control group (teaching as usual). The school with one Year 1 classroom was automatically assigned to the intervention group. As this was a whole-class (universal) programme all of the pupils in the intervention classroom received the adapted PATHS programme in place of usual SEL instruction (see the results from the interviews for SEL teaching as usual approaches). See Table 2 for the teacher characteristics and the frequency in each group, and Table 3 for the pupil characteristics and the frequency in each group.

Table 2. *Characteristics of the teachers and frequency in each group*

	Variable	Intervention (*n*)	Control (*n*)
School type	Primary	5	4
	Special needs	1	1
Staff role	Teacher	6	5
Gender	Female	5	4
	Male	1	1

Table 3. *Characteristics of the pupils and frequency in each group*

	Variable	Intervention *(n)*	Control *(n)*
School type	Primary	94	47
	Special needs	3	4
Gender	Female	42	24
	Male	55	27

Procedure

Each of the teachers (n = 6) in the intervention classes attended the same one-day training session that took place in a classroom in one of the schools in February 2019. Training was delivered by the first author and included discussion of the key theories and research, the concepts and materials relevant to the adapted PATHS programme, and included detail about the pilot study. To familiarise teachers with the materials, we conducted an activity in which teachers worked in groups to explore the lesson plans and materials and were invited to ask any questions related to implementing and conducting lessons. After the training, teachers were provided with copies of the training slides, the programme resources and implementation checklists, and copies of the teacher-report measures.

Teachers completed baseline measures for pupils in February 2019 (*n* intervention = 97, *n* control = 51). The adapted PATHS curriculum was implemented for 12 weeks from the end of March to the start of July 2019. After each lesson, self-report implementation checklists were completed. Teachers completed post-intervention measures for pupils in July (*n* intervention = 97, *n* control = 51) once implementation of PATHS had concluded. In addition, the first author completed classroom observations in May 2019 and two focus group interviews were conducted with the intervention teachers in July 2019.

Measures

We conducted assessments of social and emotional outcomes pre- and post-intervention for all pupils to provide an estimation of the effects of the adapted PATHS programme. We measured pupils' emotion regulation using the teacher respondent Emotion Regulation Checklist (ERC) (Shields & Cicchetti, 1997). The ERC is a 24-item questionnaire that assesses the frequency of behaviours using a four-point Likert scale (from 1 = never to 4 = almost always). The ERC yields two subscales: emotion regulation and lability/negativity. Emotion regulation is evaluated by the total sum of eight items describing situationally appropriate affective displays, and empathy and emotional self-awareness. Higher scores (4-

32) on the emotion regulation scale indicate a greater capacity to manage and modulate one's emotions. The lability/negativity subscale comprises total scores of 15 items that assess inflexibility, dysregulated negative affect, and unpredictability and suddenness of mood change. Higher scores (15-60) on the lability/negativity subscale indicate excessive emotional reactions and frequent mood changes that are unrelated to stimuli. The ERC has shown a high reliability of $\alpha = 0.85$ (Shields & Cicchetti, 1997).

Pupils' social and emotional behaviours were measured using the teacher version Social Competence Scale (SCS). This 25-item scale was developed for the Fast Track project (Corrigan, 2002). The scale includes items measuring prosocial behaviour/communication skills (nine items), emotional regulation skills (nine items) and academic behaviour (seven items). Teachers rate items on a five-point Likert scale (from 0 = not at all to 4 = very well) according to how well behaviour characterise the behaviour of each pupil. Higher scores on the scale (0 – 100) indicate greater social competence. The internal consistency of the total scale was .96 (Conduct Problems Prevention Research Group, 1995).

We administered the teacher report Strengths & Difficulties Questionnaire (SDQ; Goodman, 2001) to measure pupils' mental health difficulties. The instrument consists of five subscales, each with five-items, focusing on conduct problems, emotional problems, hyperactivity/ inattention, peer relationship problems and prosocial behaviours. As a result of using the SCS to measure social competencies, we chose to remove the prosocial behaviours subscale and use the four behavioural difficulties subscales, only. Teachers rated items on a three-point Likert scale (from 0 = not true to 2 = certainly true). Scores from the four behavioural subscales are combined to produce a total difficulties score ranging from 0 to 40. Total difficulties scores are categorised into three categories: normal (scores between 0 - 11), borderline (scores between 12 - 15) and abnormal (16 - 40). According to Goodman (2001), this test has a satisfactory level of internal consistency (r = .73) and test-retest stability (r = .62).

Five teachers completed the self-report implementation checklists for each completed lesson, and one teacher did not complete any. The eight-item checklists assessed the date of (1) delivery, and (2) time taken to deliver a lesson (dosage); (3) the extent that the required lesson activities were completed, (4) structure of lessons activities followed, and (5) generalisation activities were used (fidelity); (6) pupil participation, (7) interest and engagement (responsiveness); and (8) preparedness for the lesson (quality). The checklist was created based on existing rubrics used in previous PATHS studies (Faria et al., 2013; Kam et al., 2003) and a PATHS lesson observation form available from the Penn State Prevention Research Center (https://www.episcenter.psu.edu/node/605). All items except the dosage data (three-items) were completed using a four-point Likert scale. Higher scores represent greater fidelity, quality, and responsiveness. See Appendix I for an example self-report implementation checklist.

Observations

Structured observations of PATHS lessons were completed in three intervention classrooms. We were unable to arrange a date with the other three intervention classrooms. To collect data, we created an observation checklist that closely aligned with the teacher-report implementation checklist to assess the following features: the extent that the required lesson activities were completed; the structure of lessons activities followed (fidelity); how prepared, enthusiastic and interesting the delivery of the lesson was (quality); and, to what extent the pupils were interested and engaged in the lesson (responsiveness). Each item was scored using a four-point Likert scale. Higher scores represent greater fidelity, quality, and responsiveness. During the observations we also documented adaptions made to the PATHS lessons.

During the observations, the teacher conducted a PATHS lesson whilst the first author observed and completed the checklist. To assess the reliability of observations, a teacher from the adaption team (CS) attended one of the observations, and agreed with 100% of the ratings. See Appendix J for the example observation checklist.

Focus groups

We conducted two semi-structured focus group interviews. One with four teachers and the other with two teachers. The remaining teacher was unable to attend.

Focus groups were held in one of the participating schools, which lasted approximately 40-minutes. We prepared a semi-structured interview guide that included open-ended question prompts prepared by the first author. The question prompts focused on the suitability and acceptability of the adapted PATHS programme, such as perceptions of compatibility with context (e.g., needs of pupils and teachers), practical fit (e.g., ease of use, convenience, appropriateness) and teachers' knowledge and skills (e.g., understanding of SEL). We also explored the current SEL approaches used, and importantly, further adaptions that could made to the programme, including the small structural adaption to the lesson and use of additional resources that were documented during observations. Discussions were led by the first author and were recorded using a hand-held recording device, and some notes were made. Prior to commencement, the participants were informed of the topics for discussion.

Analysis

This close-to-practice-research study was part of a wider project with a school cluster and the intervention was adapted specifically for their context. Subsequently, there is no issue of 'statistical' generalisation to a larger group of schools as the outcomes are only relevant to the population being studied in this cluster of schools. Moreover, this pilot study utilised a quasi-experimental trial design in which pre-existing classrooms were placed into groups, resulting in imbalanced groups, and a considerably larger sample in the intervention group at pre- and post-test. Accordingly, and in alignment with the rationale and statistical methods used by Gorard, Siddiqui and See (2014) in their evaluation of a Switch-on Reading, we chose to calculate descriptive statistics and effect sizes, only. Specifically, the two outcomes of interest were the difference scores, which were the differences between the mean pre-test and post-test scores, and the effect sizes. The post-test scores and difference scores were averaged for each group (intervention and control), and the difference between the average scores were expressed as an ‘effect’ size. The effect size used was Hedges’ g, which was calculated by dividing the difference by the pooled standard deviation of the gain scores for both groups.

We calculated descriptive statistics to represent dosage, fidelity, quality and responsiveness of implementation for the data derived from the teacher-report implementation checklists and observations.

The data from the focus group interviews was analysed using thematic analysis, as outlined by Braun and Clarke (2006). For the thematic analysis, the recorded data was transcribed and read at least twice before salient and interesting commentary were coded and organised into initial thematic categories. We then reviewed, defined and labelled the final themes and subthemes, and presented them in a narrative description.

We adopted a hybrid approach to the process of identifying and labelling themes, which involved both a deductive and inductive approach (Fereday & Muir-Cochrane, 2006). This enabled our analysis to be informed by existing theories and evidence from the implementation literature and particularly from PATHS research (e.g., Humphrey et al., 2018; Pegram et al., under review – Chapter 4; Rogers, 2003), and ensured that any emergent themes were included.

Results

Five teachers completed the self-report implementation checklists and one teacher completed none. During the intervention period (March - July), none of the teachers completed all of the weekly lessons. Across the five teachers, the median number of lessons was 9 (75%), and the median time taken to conduct a lesson was 30 minutes. The median teacher-report fidelity score was 10 (possible scores range from 3-12), the median quality score was 4 (possible scores range from 1-4), and the median responsiveness score was 7.50 (possible scores range from 2-8), which all represent high levels of implementation. See Table 4 for the teachers' self-report implementation data.

Table 4. *Displays the average implementation scores for dosage, fidelity, quality and responsiveness, across the intended 12 weeks of lessons*

Teacher	Frequency of delivery (%)	*Mdn* delivery time taken in minutes	*Mdn* fidelity score	*Mdn* quality score	*Mdn* responsiveness score
1	NR	NR	NR	NR	NR
2	8 (67)	30	10.50	4	7.50
*3	9 (75)	60	10	3	6
4	9 (75)	20	10	3	8
8	10 (83)	32.50	11	4	8
6	11 (92)	30	10	4	6

* Special school, NR = no response

Across the three classroom observations the median fidelity score was 8 (possible scores range from 2-8), the median quality score was 4 (possible scores range from 1-4), and the median responsiveness score was 8 (possible scores range from 2-8), which all represent high levels of implementation. Table 5 displays the implementation scores from each observation.

Table 5. *Displays the dosage, fidelity, quality and responsiveness data derived from the observations*

Observation	Time taken in minutes	Fidelity score	Quality score	Responsiveness score
1	20	8	4	7
2	30	8	4	8
3	30	8	4	8

The reported effect sizes (+0.42, +0.33, +0.40 and +0.43) show medium positive impact on pupils' social and emotional competence from the adapted PATHS programme (see Tables 6 - 8). Both groups (intervention and control) had similar scores at pre-test, which suggests equivalence at baseline on all measures; indicating that the intervention was effective.

The mean score from the ERC subscale suggests that at pre-test the intervention group had moderate to high ability to manage and regulate their emotions (intervention *M* = 25.70, control *M* = 24.55), and low to moderate levels of emotion dysregulation (intervention *M* = 21.52, control *M* = 22.76). At post-test, the score gain on the emotion regulation subscale for the intervention group represents greater emotion management and regulation (*M* difference = 1.98); and, a score decrease on the lability/negativity scale (*M* difference = 1.60), suggests lower levels of emotion dysregulation. See Table 6 for the mean scores and effect sizes.

Table 6. *Results from the teacher-report pupil Emotion Regulation Checklist*

Emotion Regulation Checklist: emotion regulation subscale						
Treatment group	*N*	*M* pre-test score	*M* post-test score	Difference	*SD*	*ES*
Intervention	97	25.70	27.68	1.98	3.81	
Control	51	24.55	24.96	0.41	3.29	
Overall	148	25.30	26.74	1.44	3.70	+0.42
Emotion Regulation Checklist: lability/negativity subscale						
Treatment group	*N*	*M* pre-test score	*M* post- test	Difference	*SD*	*ES*
Intervention	97	21.52	19.92	1.60	3.99	
Control	51	22.76	22.61	0.16	4.78	
Overall	148	21.95	20.84	1.10	4.32	+0.33

The mean scores on the SCS, suggest that at pre-test pupils in the intervention group had moderate levels of social competence. At post-test, gains in scores (*M* difference 10.37) represent an increase in pupils' social competence. The relatively large standard deviation of the difference scores indicates a variation in pupils' social competence scores, which range from -28 to 57. Mean scores and the effect size are represented in Table 7.

Table 7. *Results from the teacher-report pupil Social Competence Scale (SCS)*

Social competence scale						
Treatment group	*N*	*M* pre-test score	*M* post-test score	Difference	*SD*	*ES*
Intervention	97	67.58	77.95	10.37	16.26	
Control	51	69.96	74.24	4.27	11.90	
Overall	148	68.40	76.67	8.27	15.14	+0.40

At pre-test, the mean total difficulties scores on the SDQ for the intervention group was within the normal range of 0-11). At post-test, a decrease in scores (M difference = 2.52) represents a decrease in mental health difficulties (e.g., conduct problems, emotional problems, hyperactivity/ inattention, peer relationship problems). Mean scores and effect sizes are presented in Table 8.

Table 8. *Results from the teacher-report pupil Strengths and Difficulties Questionnaire (SDQ)*

Strengths & difficulties questionnaire						
Treatment group	*N*	*M* pre-test score	*M* post-test score	Difference	*SD*	*ES*
Intervention	97	7.12	4.61	2.52	4.15	
Control	51	5.98	5.20	0.78	3.84	
Overall	148	6.73	4.81	1.92	4.11	+0.43

The following section provides a brief narrative summary, with relevant data excerpts, for each of the eight themes that we identified from our focus group interviews with the intervention teachers.

Theme 1: Approaches to SEL

Social and Emotional Aspects of Learning (SEAL)

All of the teachers were using SEAL: *"We've been following SEAL for a couple of years. It's easy to follow"*. Some of the teachers were also delivered SEAL alongside other activities: *"we've looked at SEAL, with some of our own activities."* However, comments suggest that some teachers struggled to fully embed and deliver SEAL lessons regularly*: "I was delivering the SEAL programme. We did, but it was very on and off. It was hard finding a structure to it."* "*There's no set plan for SEAL".*

Other approaches

Teachers also discussed alternative methods to promote pupil wellbeing: *"I ran nurture groups, which were in house. We did a lot of circle time and things like that within my classroom."* Finally, a couple of teachers referred to previous use of the original PATHS UK programme with mixed perceptions: *"We had used PATHS but it was a bit lengthy and a bit arduous". "There were moments of PATHS that were good".*

Theme 2: Compatible with beliefs, values and need

All of the teachers in both focus groups spoke about the importance of developing pupils' social and emotional competence, and perceived it to be a crucial part of pupils' education: *"If the skills are not developed and mastered, then pupils struggle to learn, because there's a barrier of not being able to be social or deal with their emotions... It underpins everything"*. Teachers also reflected on how important they believe the skills are for pupils outside of the classroom: *"Each year you're dealing with different situations and things get more difficult on the playground, and they (pupils) need to know how to deal with those situations".*

Many of the teachers noted the usefulness of the adapted programme for managing problematic behaviour that was displayed by some groups of pupils: *"I've noticed with the group of children I've got, some of the boys don't get on with each other, and then there is a group of girls who are in their own little group, so it's really helped with that, coping and* *managing their behaviours";* and, for meeting the needs of pupils that experience difficult home lives*"Some of them (referring to the pupils) come back to school the next day with a lot of baggage"... So it has been good for that".*

Theme 3: Acceptable and suitable

Resources and materials

The majority of comments reflected teachers' satisfaction with the programme resources and materials: *"It was easy to deliver every week…there was a script there and things like that, and it was refreshing to have some resources that were already laminated."* The resources also helped to engage pupils in the lesson: *"My children really enjoyed it. I think they liked having those visuals, those pictures. I have them hung up in class, so they knew that when I take them down it's circle time, and they sit automatically in a circle".*

Structure and content

When asked about the structure and content of the adapted PATHS programme, teachers spoke positively: *"Straight up very very easy, straightforward, and worked really really well".* Teachers indicated that delivery did not require much preparation: *"It was really easy to pick up and just do".* For many, the sequencing of lessons and progression of learning each week was viewed as advantageous: *"It was nice to pick up where you left off last week and switch to this week. I liked the structure and that it builds. Yeah, I did like that. SEAL doesn't have that, it's picky".*

Comments also suggest that teachers believed the structure and content of the programme was pitched at the right developmental level for Year 1 pupils: *"The children found it easy and straightforward to follow".* Some teachers had misconceptions that the lesson objective would be too advanced: *"I was thinking, gosh, they're never going to understand the compliment, but actually teaching it and doing it with them they've really got to grips with what a compliment is"*; and others had misconceptions that some of the lesson objectives were too simplistic: *"For some of the objectives, I thought they should know already, but not all of them did".*

Dosage and delivery time

Lastly, the recommended dosage and time required to deliver a lesson was something that teachers found conducive to the regular delivery of the adapted PATHS programme: *"Yeah, once a week, I think is ideal. I think it's plenty to be doing something like this regularly".* Furthermore, the teachers perceived the length of lesson deliver time to be more advantageous than the original PATHS programme: *"Like, they were short as well, you can find that PATHS* (referring to the original PATHS) *were going on too long".*

Theme 4: Dissatisfaction

Not all of the teachers liked how concepts were re-visited in quick succession. Teachers reported negative responses from pupils: *Happy and sad were repeated weren't they? …Even though there was elements of it that were different, mine said, Oh not happy and sad again".* Although the minority, one teacher felt the feeling faces resources did not always display accurate emotions: "*The resources need to be clearer. So for some of the expressions, I may have looked at it and thought there was something else."*

Theme 5: Negative impact from competing priorities

Teachers also reported that the most significant barrier to conducting all of the 12 lessons was finding the time to fit the lessons into an already full timetable. Comments suggest that competing priorities and other curriculum demands often took priority and meant that lesson delivery time was reduced and/or lessons were missed: *"I didn't quite get around to doing the final lesson, but did everything else… we had sports day practices, just lots of things like that going on, we've had healthy living week. But yeah, all these things put pressure on a tight timetable". "So if there's anything that's got to go, unfortunately, it has to be those (referring to adapted PATHS) lessons that you're less accountable for".*

Theme 6: Further adaptions

Use of additional resources

Throughout the discussions, teachers made clear that they liked the structure and content of the programme and implemented these elements with fidelity. However, it was suggested that additional resources were required to increase the suitability of lessons for pupils. Subsequently, teachers made small modifications to lessons by including additional activities and resources to make lessons more fun: *"I followed it very rigidly… we added our own sort of spark to it, really. We looked for other additional role-play opportunities, or other ways to make it a little bit more fun as well for them to them to get the best out of it. We found some YouTube songs and clips that fitted in really nicely". "Even though we followed the structure, we still put our little own little twist on it."*

Minor structural modifications

One teacher referred to retaining the content of the programme and modifying the structure of the first lesson (to teach classroom expectations/rules), to accommodate pupil need: *"I did one rule a week, instead of all at once. I just focused on good sit this week. And then next week I introduced a good listener…"* Similarly, another teacher chose to split her class into two groups and deliver the lesson twice:*"I split mine into two sessions, so it wasn't with the whole class. So had a lesson of 32 and a lesson of 15, which made it much more intimate… So I did it twice, which helped me as well".*

Theme 7: Intentions to sustain use

When teachers were asked whether they would like to continue using the adapted PATHS curriculum to teach whole class social and emotional skills, they all responded positively and suggested that they would like to use the programme from the beginning of September, with their new cohort of pupils: *"I find it invaluable, really valuable. I definitely want to use it in the first term and expand on, you know, use the criteria that you've set out, and also use a few extra things that we've incorporated into it along the way to make it a* *little bit more on their level". "Definitely. I would like to do it, but perhaps in September for the new class".*

Theme 8: Negative impact from competing programmes

There were some concerns about how the adapted programme was going to fit in with existing SEL programmes, which was predominately SEAL: *"Be nice to see how maybe it could either compliment SEAL, because the rest of the classes are using SEAL." "Are we going to go back to SEAL, or is it instead of? I want to use it again. But I also need to consider how it is going to fit in with what the rest of the school is doing".*

Discussion

In this study we worked with teachers from a cluster of schools in Wales to culturally and contextually adapt the PATHS programme. We then evaluated if the adapted programme would lead to positive attitudes, high levels of implementation, intentions to sustain use, and positive impact on pupils' social-emotional competencies. Our results revealed that the adaption objectives were achieved, teachers' attitudes towards the adapted PATHS programme were positive, levels of implementation were high, and evidence of positive medium effects on pupils' social-emotional skills were observed. Thematic analysis identified eight themes: (1) approaches to SEL, (2) compatible with beliefs, values and need, (3) acceptable and suitable, (4) dissatisfaction, (5) negative impact from competing priorities, (6) further adaptions (7) intentions to sustain use, and (8) negative impact from competing programmes.

Before participating in this study, the intervention teachers revealed that they used SEAL to teach social and emotional skills (theme 1). However, comments suggested that some teachers struggled to deliver SEAL frequently and consistently due to a lack of structure in the programme. Two teachers referred to using PATHS, and the group agreed with a comment from one teacher that *"it was lengthy and arduous"*, which aligns with the issues reported in other PATHS implementation research (e.g., Faria et al., 2013; Humphreys et al., 2018).

There is limited research that describes the process of adapting PATHS and only one previous study (Inam et al. 2015) has utilised mixed-methods pilot research to evaluate a newly adapted programme. To advance the evidence-base, we aimed to systematically describe the process of adapting the PATHS UK Reception/Year 1 programme and conduct a feasibility evaluation of the adapted programme. Primarily, we sought to understand if a more culturally and contextually relevant programme would lead to more positive teacher attitudes, increased levels of implementation, and intentions to sustain use. In doing so, we hoped to increase the use of evidence-based programmes in the cluster schools, and contribute to the literature on the best ways to increase adoption and sustained implementation of evidence-based provision in schools.

We described the process of systematic adaption of PATHS® using the heuristic framework described by Barrera & Castro (2006). One of the core recommendations in the model and across the adaption literature is the involvement of different stakeholders and community members with a vested interest in the adaptation process. The role of the teachers in this study can be viewed as a critically important aspect.

In accordance with the views of Lau (2006), we adopted a 'selective and directed' approach to the gathering of information and only focused on the information relevant to the objectives of adaption. According to Lau (2006), indiscriminate adaptions might lead to adjustments that later prove to be unsuitable. Similar to Inam et al. (2015), we conducted several group meetings with teachers with expertise in school wellbeing to discussion and evaluate the issues and objectives. In addition to, and alignment with the guidance by Barrera and Castro (2006), we also chose to evaluate the literature for other evidence-based kernels to teach social-emotional skills, and for dosage recommendations. As described by Barrera and Castro (2006), and based on our primary aims of the study, we chose to conduct a small 12-week pilot study to determine the likely impact of the adapted programme on pupil outcomes. We also conducted in depth interviews to determine if the objectives of adaption had been achieved, the acceptability and suitability of the programme, and to identify issues and other adaptions. In addition to the guidance, we chose to collect both teacher-report and observation implementation data to assess multiple components of implementation, recommended by Durlak and DuPre (2008) to gain a more thorough understanding of the extent that PATHS was implemented. Lastly, for the adaption refinement stage in the framework, we planned to present the data from the evaluation to the group and discuss further adaptations to the programme. Unfortunately, the COVID-19 pandemic interrupted the final adaption refinement phase of the study.

Greater adoption and implementation of evidence-based programmes in schools is likely if adopters perceive programmes to be compatible with context (Pegram et al., in review – Chapter 4; Rogers, 2003). We had three main adaption objectives for this study, which were to: (1) sufficiently anglicise the programme content and materials and include opportunities to teach Welsh; (2) condense the content whilst retaining the key kernels (core ingredients) of the programme; and (3) reduce the requirement of twice-weekly dosage. We believe these adaptation aims were achieved. We retained much of the core content and format of the PATHS® programme, including the generalisation techniques and activities to be implemented throughout the school day. We rewrote the content and condensed the information into lesson plans, scripts and supporting materials, and adapted the lesson activities and ensured all language, names and cultural references were appropriately anglicised. We also incorporated opportunities to practice the Welsh language throughout the script and activities. We created and included a list of discretionary extension activities so teachers could personalise teaching, and developed a host of supporting materials to replace all of the PATHS® materials (e.g., feeling faces), which were all laminated. Finally, we recommended that PATHS lessons be delivered once a week for approximately 30-40 minutes.

The data derived from the focus groups revealed that both the objectives of adaption were achieved, and that overall teachers' attitudes towards the adapted programme were positive. Teachers perceived the resources and structure of lessons to be acceptable and suitable (theme 3). Throughout the interviews, teachers spoke positively about the lesson plans and resources, and described the structure of the lessons as straight-forward, and easy to follow, and the content as developmentally appropriate. Furthermore, teachers indicated that lessons did not require much preparation. Most teachers liked the sequencing of lessons and progression of learning each week, and viewed this as advantageous when compared to Social Emotional Aspects of Learning (SEAL).

Comments from teachers also revealed that the adapted programme was compatible with teachers' values, beliefs, and needs (theme 2). Teachers remarked on the importance of the social and emotional skills that were taught through PATHS, and claimed that the skills *"underpinned"* pupils' education. Furthermore, the programme was considered a good fit with some of the more specific behavioural needs of pupils that teachers often deal with.

Notably, all the teachers reported that the suggested dosage (one session per week) and time required to deliver the PATHS lessons was conducive to the regular delivery of the adapted PATHS programme. A common barrier to regular implementation of programmes, commonly reported throughout PATHS and other implementation literature (e.g., Faria et al., 2018; Humphreys et al., 2016), is the influence of competing priorities (theme 5) such as other curriculum demands and activities, and this is a feature in this study. Comments suggest that lessons were missed and/or delivery time was reduced as a result of prioritising other academic subjects and activities such as sports days and healthy living week. Evidence from Pegram et al. (2022 – Chapter 3) and Chapter 5 indicates that schools in this cluster are already using a large number of programmes and strategies, and this that might hinder the uptake and implementation of an additional programme.

Some teachers did express some dissatisfaction with the adapted programme (theme 4). One teacher did not like the repetition of concepts in two of the lessons and considered them to be in quick succession. However, as this was a 12-week programme the lessons with revisited concepts would be adequately spread throughout a full school year and this was not considered to be problematic. Another teacher revealed some dissatisfaction with the accuracy of the feeling faces resource.

In alignment with Humphrey et al. (2016, p.23), we also found that teachers made predominately 'surface level' adaptions to the programme (Theme 6). For example, teachers used additional resources to personalise lessons, and modified structure to accommodate pupil need. Based on these results, we can conclude that further adaptions refinements are not required.

Evidence from the quasi-experimental trial found medium positive effects of the adapted PATHS programme on pupils' social and emotional competencies when delivered once a week across a 12-week intervention period. Specifically, the adapted PATHS programme had medium positive causal effect sizes on pupils' emotion regulation (emotion regulation subscale +0.42, emotional lability/negativity subscale +0.33), pupils social skills (+0.40) and pupils behavioural difficulties (+0.43) (Hedges, 2008). When using Kraft's (2020) revised benchmarks for interpreting research outcomes for educational interventions for preschool to Key Stage 5, these effect sizes would be considered large. However, it is important to note that we would expect to see larger causal effects from the SEL programme on outcomes that are directly related (e.g., pupils' social-emotional skills), particularly when the skills are measured soon after the intervention ended, and particularly with younger learners. It is also worth acknowledging that despite a predominately healthy sample of pupils (e.g., on average pupils had moderate-high levels of social and emotional competence and were within healthy ranges on the SDQ, at baseline), moderate-large effects were observed. Furthermore, we suggest that our results are promising considering that Kraft (2020) reports that universal interventions typically produce more modest effects than interventions targeted at those that need it the most.

The findings from teacher-report checklists and observations suggest that the adapted programme did lead to adoption and high levels of implementation. Although none of the teachers completed all 12 lessons, on average teachers did deliver nine lessons (75%), and the mean delivery time was 30-minutes. Furthermore, the teacher-report implementation checklist and observations revealed that the adapted programme was delivered with high levels of fidelity, quality and responsiveness. The frequency of lesson delivery has been described as a core component of implementation (Durlak & DuPre, 2008). Humphrey et al. (2016, 2018) suggests that less than optimal dosage reduces programme effectiveness. However, according to Humphrey et al. (2016), a specific dose-response relationship for PATHS had not been established. Conversely, a number of studies that have collected dosage data, found that on average lesson delivery was approximately 50 per cent, and despite less than optimal delivery, promising findings were reported (e.g., Berry et al., 2015, Faria et al., 2013; Humphrey et al., 2016; Kam et al., 2011). Moreover, this finding is consistent across other implementation literature (Durlak & DuPre, 2008). Similarly, we found that despite less than optimal dosage, medium-large effect sizes were reported.

Based on feedback from teachers in the adaption team and the aforementioned evidence, we opted for a once-weekly dose of the adapted PATHS programme, and found positive effects on pupils' social and emotional outcomes from the adapted programme, when approximately 75 per cent of lessons were delivered, which was on average nine times across a 12-week intervention period. Evidence such as this has led researchers (e.g., Faria et al., 2013) to argue that implementation quality may be a greater predictor of effectiveness than dosage, which might explain, in part, the large positive effects found in this study. Furthermore, Humphrey et al. (2018) demonstrated that levels of implementation quality and responsiveness were significantly associated with pupils' social-emotional skill acquisition.

Finally, when we asked teachers if they would consider continuing to use the adapted programme, many revealed that they would like to implement it with a new cohort at the beginning of an academic year. However, concerns were raised about how the adapted programme was going to fit in with SEAL (theme 8), which suggests a lack of capacity to sustain use of an evidence-based programme.

Limitations

Unlike the cultural adaption of PATHS conducted by Inam et al. (2015), we did not include or consult with the programme developers, which may have resulted in some loss of core components or structure essential to programme efficacy. We also believe it is important to highlight that the PATHS adaption team consisted of the first author and the cluster teachers, and the first author planned and conducted the evaluation, as such there was vested interest in the success of the programme, which means elements of bias cannot not be ruled out. Furthermore, we adapted the PATHS programme specifically for the participating schools, staff and pupils and, therefore, the programme is not intended to be generalised to other schools. The main aim of this study was to evaluate school staff attitudes towards the adapted programme, and the extent to which the adapted PATHS programme was implemented. However, the underlying process and principles of adaption is intended to inform policy and research.

For the pilot trial of the adapted programme, it was not possible to randomly assign participants to intervention and control groups, which means we cannot assume that the adapted PATHS programme contributed to the medium-large effect sizes. As a result of non-random assignment, there was a large imbalance in groups, and this may have affected results. However, the results from the pre-test data suggests that there was equivalence between groups on all measures at pre-test. To mitigate the effects from unequal groups, we calculated Hedge's g using the pooled standard deviation of the gain scores (Hedges, 2008). Notably, to measure pupils’ social-emotional competencies, we used non-standardised measures, which Wolf and Harbatkin (2022) suggest often produce larger effects sizes. They also report that effect sizes are generally higher in younger populations. Whilst the medium-large effect sizes reported during the 12-week intervention phase are promising, we should also note that a plausible explanation for the increases in pupils’ social-emotional competencies may be a result of normal developmental/maturation.

Another limitation pertains to the measures of implementation. To mitigate the bias from teacher-report ratings of implementation, we conducted observations to triangulate with teacher-report data, and a good fit was identified. Although generally considered to be a more valid measure of implementation than teacher self-report surveys, a single independent observation provides only a snapshot of implementation activity that may or may not be representative of implementation across the 12 weeks of intervention delivery. Humphrey et al. (2016) recommends multiple observations of implementation to improve reliability and identified any patterns in the data. However, for this study limited resources precluded multiple observations. Furthermore, we are also unable to rule out the influence of ‘observer effects’.

We acknowledge that future research using a more robust, experimental design that includes cluster randomisation at the level of the classroom, and cluster level analysis, is now required. Furthermore, we also recognise the need to collect other useful information such as pupil gender and free school meal information, as variable effects have been reported on these subset of pupils (Humphreys et al., 2018). And, a more robust evaluation will evaluate the longer term effects from the adapted programme, which currently are unknown.

Implications

Based on the positive findings from this study, we now recommend that a robust trial of the programme is conducted to evaluate the effects of the programme across a full academic year.

The process of adaption detailed in this study provides guidelines for future researchers interested in systematically adapting PATHS and other school-based programmes, using the Heuristic framework (Barrera & Castro, 2006).

We believe our results have highlighted the importance of careful adaption of programmes for implementation research. The adaption of a programme is often viewed as a failure to implement a programme with fidelity. When evaluated, most research focuses on measuring the extent that a programme has been implemented (fidelity and dosage) and how well (quality) a programme has been implemented, and describes the data using arbitrary statistics (Durlak & DuPre, 2008; Humphreys et al., 2018). We argue that adaptions should also be evaluated. The findings from Chapter 4 and 5, and Rogers (2003), suggests that school staff are more likely to adopt evidence-based programmes if they are perceived as compatible with context, and a programme that provides clear guidance for implementation fidelity and adaption is more likely to fit a broader variety of contexts. If teachers are to adopt and implement a programme with higher levels of fidelity and dosage, then it is important that they have information about what core components of a programme should be retained to ensure effectiveness, and the components that can be adapted. In alignment with the evidence building framework developed by (Owen et al., 2022), we propose that greater efforts are required to more robustly field test programmes or teaching approaches during phase 1 (phase 1: define, manualise and test interventions), before moving to efficacy testing.

Conclusion

This study systematically described the process of adapting the PATHS® UK reception/Year 1 programme using the heuristic framework by Barrera and Castro (2006). The results from the evaluation of the newly adapted PATHS programme suggest that, the objectives of adaption were achieved, teachers' attitudes towards the PATHS programme were positive, levels of implementation were high, and medium-large effects on pupils' social and emotional outcomes were observed. We found that the biggest threat to optimal delivery of the programme was the impact of time and competing curriculum demands. We also found that teachers are now more likely to consider using the adapted PATHS programme; however, competing programmes may influence decisions to sustain use. Now, we recommend that robust research is conducted to evaluate the efficacy of the programme when used across a full academic year. The process of adaptation in this study highlights the need for more rigorous process evaluations and feasibility research prior to commencing efficacy and effectiveness trials.

Chapter 7: Discussion of thesis

Discussion

Broad Overview

To improve the quality and equity of the education system in Wales and across the UK, there has been a renewed focus on creating a more evidence-based or evidence-informed practice (Donaldson, 2015; DfE, 2010, 2016; Institute for Effective Education, 2019; OECD, 2014; 2017; Welsh Government, 2014; 2021). An evidence-informed approach includes the use of the 'best' available evidence from worldwide research, the professional judgement and experience of educators, and system level and classroom data, to improve practice (Gorard, See, & Siddiqui, 2020; Nelson & Campbell, 2017; Sharples, 2013). Within a self-improving system, educators are now expected to develop methods to improve the quality of teaching practice through teacher enquiry and the use of external research through collaborative working and sharing best practice methods (Greany, 2015; OECD, 2014; Slavin, 2020; Welsh Government, 2017; 2021). However, there is still an inconsistent use of evidence to inform teaching practice (Brown & Zhang, 2016; Coldwell et al., 2017) and the ways in which evidence-into-use can be achieved are not explicitly defined in the literature (Gorard et al., 2020; Godfrey & Brown, 2018). Efforts to increase the use evidence have been identified in the literature, such as providing teacher-friendly summaries of evidence, and in-service training (Gorard et al., 2020; Hemsley-Brown & Sharp, 2003; Nelson & O'Beirne, 2014; Slavin, 2019).

Following recommendations from the OECD (2014), the Welsh Government embarked on a series of large-scale policy initiatives to improve the quality of education, which included developing a more collaborative and research-informed self-improving system: see *Qualified for Life: An Education Improvement Plan* (Welsh Government, 2014), and *Education in Wales: Our National Mission* (Welsh Government, 2017). The research on which this thesis is based was conducted with a school cluster in Wales. The cluster had entered into a collaborative research partnership with regional consortia and Bangor University to co-construct close-to-practice research that aligns with the values and priorities of the organisations involved, and importantly addresses the issues pertinent to the schools (OECD, 2014; Welsh Government, 2014, 2017). The primary aim for the partnership, was to increase the use of evidence-based methods in the cluster schools and move them towards a more evidence-informed practice. A priority and request from the cluster senior leaders was to review the evidence base for the interventions already in use across schools. An additional priority was to focus the research on improving the quality of well-being provision. The main research aims set out by GwE and CIEREI, Bangor University, were to explore ways to increase the use of evidence-based methods in schools.

Chapter 3 evaluates the range of interventions used in the schools and assesses the evidence base for the interventions. A follow-up study in Chapter 3 investigates the impact of teacher-friendly reports of evidence on existing provision. Chapter 4 explores the decision-making process of senior leaders when choosing academic and wellbeing programmes. Chapter 5 explores the extent that teachers implemented the GBG following typical in-service training, and the factors that facilitate decision to adopt or reject the procedure. Chapter 6 describes the process of adapting PATHS and includes a feasibility evaluation of the adapted programme. The evidence derived from this research aims to increase the use of evidence-based methods in the cluster, to move them towards a more evidence-informed practice. The evidence also contributes to the literature on the best ways to increase the use of evidence-based provision in schools.

Chapter 3

Very little is known about the range of intervention programmes and approaches used in UK schools, or the evidence supporting them. There is some international evidence that suggests schools do not adopt evidence-based interventions (Hallfors & Godette, 2002;

Gottfredson & Gottfredson, 2002). Evidence suggests that when choosing school provision, educators' decisions are more often informed by own experience of what works, and the experiences of teachers within school and in other schools (Nelson et al., 2017; Walker et al., 2019). The primary aim of this chapter was to explore the range of interventions used in the cluster of schools, and to assess the evidence base for the interventions, in particular to help identify the programmes or methods that possess evidence of positive causal impact on pupil outcomes. We believe that the findings from this study provide important new information on the range and evidence-base of interventions in use in a cluster of schools, and the results may be used by the cluster to inform school provision. A secondary aim of this study was to investigate whether increased knowledge about the evidence for the interventions used, presented through a summary report to school leaders, had an impact on existing provision. The results contribute to the developing evidence-base on the best way to increase the use of evidence in schools.

Our results found that one-hundred and thirty-eight interventions were in use across the 10 schools. There were 55 interventions in the cognition and learning category, 64 interventions in the social, emotional, and mental health category (and a little under half of these were programmes to teach social skills and emotional resilience) and 19 interventions in the communication and interaction category. The results from the systematic review of the evidence-base found that only 30 per cent (42) of interventions used across the cluster had causal evidence of impact on a pupil outcome. However, of these, only 11 per cent (15) had evidence of effectiveness that was rated moderate to high quality and considered promising, and 20 per cent (27) had preliminary evidence, rated low quality. Sixty-seven per cent (92) of all interventions used across the cluster had no published evidence at the time of this review. The remaining 3 per cent of interventions were shown to be ineffective. Based on data received from four schools, we found that three of the schools (two primary and one special needs school) made no changes to their provision and continued to use the same interventions that were identified during the initial screening process. The remaining primary school did made changes to their provision and discontinued use of three interventions (two had no evidence, and one had evidence of ineffectiveness).

This is the first study to empirical report the range and evidence base for interventions used in a cluster of schools, and our results suggest that schools adopt a large number of interventions that have a very limited evidence-base, with very few having evidence for a positive causal-impact on learner outcomes. Our results strongly suggest that teacher time, school budgets and implementation levels (e.g., fidelity and quality of implementation) may be negatively impacted by this large number of interventions. Previous research has reported barriers to the use of evidence that may prevent teachers from drawing on the external evidence base to inform practice, and these include a lack of time, underdeveloped research literacy skills; and a lack of relevant, accessible, and usable evidence (Cain, 2016; Gorard, 2020; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018; Walker et al., 2019). We believe that our findings indicate that these barriers continue to persist.

The evidence from our follow-up study suggests that the use of educator-friendly reviews of research that clearly and concisely present evidence for specific programme and methods may increase the use of evidence-based methods for some schools, which supports recommendations by Gorard et al. (2020), Nelson and O’Beirne (2014), and Slavin (2019). However, our results support findings of Gorard et al. (2020) and suggest that when presented with summary reports of evidence, not all schools will use this information to make informed decisions.

Chapter 4

Since the creation of the research repositories such as the EEF Toolkit, more teachers and senior leaders now report engaging in the external research evidence to inform provision

(Brown & Greany, 2018; EEF, 2018). Whilst the literature does indicate engagement with evidence, it does not demonstrate that decision-making, and importantly the provision that schools adopt into use, is underpinned by research evidence (Higgins, 2016). Evidence derived internationally (Hallfors & Godette, 2002; Gottfredson & Gottfredson, 2002), and the findings from Chapter 3 suggest that much of school provision is not informed by research evidence. Evidence derived through survey research suggests that when choosing school provision, educators' decisions are more often informed by own experience of what works, and the experiences of educators within school and in other schools (Nelson et al., 2017; Walker et al., 2019). The aim of the study reported in Chapter 4 was to explore in detail, using semi-structured interviews, the factors that inform senior leaders' decisions when choosing academic and wellbeing programmes. To gain a realistic insight into the contextual factors that inform decision-making, we chose to focus the interview questions to specific programmes that the head teachers had recently adopted. The findings from the study provide insight into the best way to increase the uptake of evidence-based programmes and practices and may add further detail to the literature. Using thematic analysis, we identified three main themes when discussing the selection of both wellbeing and academic programmes: (1) experience and advice from education professionals, (2) compatibility, (3) sales and marketing. We also identified a fourth theme (4) accessing the evidence base, although this was only applicable when discussing academic programmes.

The themes derived from our data support the findings derived from previous survey-research (Nelson et al., 2017; Walker et al., 2019) and builds on this evidence base. Like the findings from Nelson et al. (2017) and Walker et al. (2019), and in alignment with diffusions of innovations theory (Rogers, 2003), we found that the advice and experience from staff in other schools, particularly other senior leaders, was frequently sought and highly trusted and valued. Rogers (2003) reports that the perceived compatibility of an innovation facilitates decisions to adopt innovations, and we also discovered that it was common practice for senior leaders to visit other schools to observe a programme of interest in use to determine the potential compatibility of the programme with context, and to understand how it functions. Importantly, we found that information gathered from other schools often influenced decisions to adopt or reject programmes. In addition to Nelson et al. (2017), Walker et al. (2019), and in alignment with Rogers (200), we found that senior leaders would seek recommendations about evidence-based methods from their local school improvement advisor (employed by GwE), who was viewed as a trusted source of knowledge for research and evidence-based practices.

In accordance with Rogers (2003) theory, we found that the perceived compatibility of programmes was a critical factor in the adoption-decision process, and that often head teachers would seek programmes that placed minimal demand on resources (staff time and money), were time saving, they could be delivered by computers or digital platforms, and could be accessed at home. Previous survey research by Walker et al. (2019) reported that teachers' decisions were also influenced by continuing non-research-based professional development events, and our findings offer some support for this as we found that some senior leaders had adopted a programme following attendance at a course. In addition, we revealed that an effective sales pitch at a training course, or effective website marketed sometimes led to the adoption of wellbeing provision. Lastly, two senior leaders referred to choosing academic programmes because they had an established evidence-base demonstrating impact on pupil outcomes, which supports Brown and Zhang (2016) and Coldwell et al.'s (2017) findings that some teachers are now using evidence to inform practice, whereas others do not routinely access research findings to inform decision-making. Overall, the findings from this study suggest that the external evidence is infrequently used to inform the programmes used in school and, in alignment with the existing literature, we suggest the external evidence repositories still currently lack relevant, practical, and useable summaries of evidence for teachers (Connelly et al., 2018; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018).

Chapter 5

Brown & Zhang (2016) and Coldwell et al. (2017) indicate that there is a lack of consistent use of evidence to inform teaching practice. Based on findings from Chapter 3 and Chapter 4 we also found that research evidence is infrequently used to inform the programmes that schools use.

The GBG is a promising strategy that has a wealth of evidence supporting its effective use in the classroom (Flowers et al., 2014). Nevertheless, like many evidence-based programmes, research suggests that the GBG has failed to translate into every-day use in the classroom (Gottfredson & Gottfredson, 2002). The concepts from Rogers (2003) diffusion of innovations theory, and our findings from Chapter 4, suggest that the perceived characteristics of the GBG (e.g., relative advantage, complexity compatibility, observability, and trial-ability) may facilitate adoption-decisions. Empirical evidence by Dijkman, Harting and van der Wal (2015) found that the compatibility of the GBG with existing teaching methods and needs facilitated decisions to adopt, and the presence of competing programmes influenced decisions to reject its use.

Research has identified the need for in-service training to be used to improve teachers' skills and knowledge of evidence-based methods (Hemsley-Brown & Sharp, 2003; Nelson & O'Beirne, 2014). Durlak and DuPre (2008), Klingner et al. (2013) and Owen et al. (2022) suggest that teachers are more likely to adopt evidence-based programmes and put them into use if trained and provided with follow-up support/coaching. With the aim of increasing the use of evidence-based provision in a cluster of schools, we trained teachers in the use of the GBG during a typical in-service training day in February. The study outlined in

Chapter 5 aimed to explore the extent that teachers would implement the GBG following typical in-service training, and the factors and perceived characteristic of the GBG that facilitate adoption-decisions. The findings from this mixed methods study contribute to the knowledge base on how best to increase the use of evidence-based methods in schools.

Our findings suggest that after completing in-service training (February) all of the teachers had good knowledge of the aims of the GBG and how to use it. We also found that six teachers that played the GBG in the classroom, viewed it as an acceptable procedure. However, only two teachers chose to adopt the GBG, three teachers rejected it immediately after training, four teachers trialled it and then chose to reject (known as an active rejection), and for two teachers their decisions were unknown. Furthermore, one teacher discontinued use of the GBG at time-point 2 (March and April), and by time-point 3 (July) none of the teachers continued to implement the GBG. Observation data from one adopting teacher at time-point two, revealed that implementation fidelity was high. However, interview data revealed that the same teacher played the game infrequently, which is commonly found across much of the implementation literature (see Durlak & DuPre, 2008; Humphrey et al., 2018).

Evidence suggests that in-service training should be used to increase teachers' skills and knowledge of evidence-based methods (Bell & Kozlowski, 2008; Hemsley-Brown & Sharp, 2003; Nelson & O'Beirne, 2014). We found that in-service training did lead to increased knowledge of the GBG and how to use it. However, our findings revealed that overall typical in-service training in the use of the GBG did not lead to the adoption of the procedure by most teachers. This suggests that in-service training may not be an effective method for increasing the use of evidence-based methods in schools. For effective implementation to take place, research indicates that re-training and coaching is required to continue to develop teachers' skills and self-efficacy (Durlak & DuPre, 2008; Greenwood et al., 2003). For example, according to Joyce and Showers (2002) only five per cent of teachers that attended training put newly learnt strategies into practice, but when re-training and coaching were included, 95 per cent of teachers implemented strategies.

We aligned our interview data with the concepts of Rogers' (2003) diffusion of innovations theory framework and identified three perceived characteristics of the GBG that facilitated decisions to reject the GBG: (1) *relative advantage*, (2) *compatibility*, and (3) *observability*. The most pertinent data revealed that four teachers rejected the GBG because they perceived other competing classroom management strategies to be more advantageous, or too similar to warrant change (*relative advantage*), which aligns with the findings from GBG implementation research (Dijkman et al., 2014; Humphrey et al., 2018). Again, in accordance with the GBG literature, three teachers rejected the procedure because they viewed it as incompatible with class needs (*compatibility*) (Humphreys et al., 2018); two teachers felt it was too advanced for Year 1 pupils, and one felt that it was superfluous to need. The most pertinent compatibility issues for four of the teachers was the time of year in which training was conducted. Teachers felt that the GBG should be introduced at the beginning of a school year, preferably in September. Lastly, the lack of opportunity to observe the GBG (*observability*) was mentioned by two teachers. Reflecting the findings of Durlak and DuPre (2008) and Greenwood et al. (2003), we suggest a lack of mastery and subsequent self-efficacy in delivering a new procedure may have facilitated rejection-decisions.

The findings from this study support previous research and reveal that in-service training alone may not be sufficient to prompt the adoption of evidence-based methods. However, we also revealed that many teachers perceived their current class management procedures to be as effective as, or too similar to, the GBG to warrant change. Furthermore, many teachers revealed that it was an inappropriate time of year to change a classroom management procedure. Evidence from Pegram et al. (2022), and reported in Chapter 3 of this thesis, indicates that the schools use a large number of programmes and strategies. As a consequence, we suggest that our findings may also reflect a lack of capacity to adopt and use yet another procedure.

Chapter 6

Adoption and implementation of programmes in schools is complex, and often evidence-based programmes perceived as incompatible are rejected by schools, or implemented with low levels of fidelity, quality and dosage (Chapter 5; Durlak & DuPre, 2008; Rogers, 2003). Cain (2016) suggests that teachers often view their context as highly complex and disparate to other school contexts and, therefore, it is unlikely that a programme will be deemed fully compatible by some teachers. According to Durlak & DuPre (2008), programmes are never implemented as intended all the time, and modifications are inevitable. Dependant on the extent of modification adaptions to a programme and less than optimal delivery can result in reduced impact (Domitrovich & Greenberg, 2000; Durlak & DuPre, 2008; Durlak et al., 2011). The PATHS programme has a sound theoretical base and a wealth of supporting evidence demonstrating its impact on pupil outcomes. However, reports suggest that implementation is variable, particularly dosage, and some teachers perceive the language and cultural references, and the length of lessons for the UK version to be incompatible with context (Faria et al., 2013; Humphrey et al., 2016, 2018). For this study we aimed to advance the evidence base by describing the process of culturally and contextually adapting PATHS for the cluster of schools. PATHS has been adapted for many cultures, but the process of adaption is rarely described, we found one study that explained the process in detail (Inam, Tariq & Zaman, 2015). In addition, through feasibility research we aimed to evaluate, if the adapted PATHS programme would lead to adoption, high levels of implementation, positive attitudes, intentions to sustain use, and positive impact on pupils' social and emotional competencies. In doing so, we hoped to increase the use of evidence-based programmes in the cluster schools and contribute to the literature on the best ways to increase adoption and levels of implementation of evidence-based methods in schools.

Our results revealed that the adaption objectives were achieved. Teachers' attitudes towards the adapted PATHS programme were overall positive, levels of implementation were high, and evidence of medium-large causal-effects on pupil social-emotional competencies were observed. Thematic analysis identified eight themes: (1) approaches to SEL; (2) compatible with beliefs, values and need; (3) acceptable and suitable; (4) dissatisfaction; (5) negative impact from competing priorities; (6) further adaptions; (7) intentions to sustain use; and (8) negative impact from competing programmes.

We had three main adaption objectives for this study, which were to: (1) sufficiently anglicise the programme content and materials and include opportunities to practise Welsh language skills; (2) condense the content whilst retaining the key kernels (core ingredients) of the programme; and (3) reduce the requirement of twice-weekly dosage. Data from our thematic analysis reflected positive attitudes towards adapted PATHS, which suggests that the adaption objectives were achieved. For example, teachers perceived the resources and structure of lessons to be acceptable and suitable (theme 2 and 3) and considered the dosage of once weekly to be conducive to the regular delivery of the adapted PATHS programme. Evidence suggests that greater adoption and implementation of evidence-based programmes in schools is likely if adopters perceive programmes to be compatible with context (Durlak & DuPre, 2008; Pegram et al., in review – Chapter 4; Rogers, 2003). Through teacher-reports of implementation and observations we found that teachers did implement adapted PATHS with high levels of fidelity, quality and dosage. However, in line with other implementation research, we found that optimal dosage was still not achieved (Berry et al., 2016; Faria et al., 2013; Humphrey, Barlow & Lendrum, 2018; Humphrey et al., 2016; Panayiotou, Humphrey

& Hennessey, 2020). Evidence from the PATHS literature and other implementation research (e.g., Faria et al., 2018; Humphreys et al., 2016) indicates that the influence of competing priorities and other curriculum demands and activities influence the regularity of delivery– this was also reported in our study (theme 5 and 8). Evidence from Pegram et al. (2022) and Chapter 5 indicates that schools in this cluster are using a large number of programmes and strategies, which may restrict teachers' capacity to implement a programme at the required levels. Furthermore, teachers also revealed that they would like to sustain use of adapted PATHS, although concerns were raised about how PATHS would fit into the curriculum with SEAL (theme 7).

Despite less than optimal delivery, evidence from the quasi-experimental trial found medium effects of the adapted PATHS programme on pupils' social and emotional competencies, when delivered once a week across a 12-week intervention period. Specifically, the adapted PATHS programme had medium positive causal effect on pupils' emotion regulation (+0.42 effect size for the emotion regulation subscale, and +0.33 effect size for the emotional lability/negativity subscale), pupils' social skills (+0.40 effect size) and pupils' behavioural difficulties (+0.43 effect size) (Hedges, 2008). However, when using Kraft's (2020) revised benchmarks for interpreting research outcomes for educational interventions for preschool to Key Stage 5, these effect sizes would be considered large. However, it is important to note that we would expect to see larger causal effects from the SEL programme on outcomes that are directly related (e.g., pupils' social and emotional skills), particularly when the skills are measured soon after the intervention ended. It is also worth acknowledging that despite a predominately healthy sample of pupils (e.g., on average pupils had moderate-high levels of social and emotional competence and were within healthy ranges on the SDQ, at baseline), moderate/large effects were observed. Furthermore, we suggest that our results are promising considering that Kraft (2020) reports that universal interventions typically produce more modest effects than interventions targeted at those that need it the most. It appears that the adapted PATHS programme is a promising option for teaching social and emotional skills to pupils in this cluster of schools.

Overall, the findings here suggest that it is possible to achieve good outcomes with evidence-based programmes that are perceived as compatible with context. Such programmes are also more likely to be implemented with higher levels of implementation and sustained over time.

Implications

Throughout each chapter we have discussed in detail several implications to address our main research aims. Below we aim to summarise the implications of this research for the cluster, policy and research.

Chapter 3 (Pegram et al., 2022) provides important new evidence, which suggests that the schools in this cluster have adopted a very high number of interventions. It is likely that this has a negative impact on school budgets and teacher efficiency, levels of implementation fidelity, quality and dosage, and the ability to develop mastery in the skills required to deliver a programme with higher levels of fidelity and quality. The interview findings from Chapters 5 and 6 also indicate that there is a lack of capacity within schools to take on evidence-based programmes and procedures because of competing priorities and programmes. We suggest that schools should evaluate and refine the range of interventions in use to maximize their impact, and to allow for the adoption of other evidence-based methods that may be more cost-effective. The cluster schools might consider using a process from health economics such as Programme Budgeting Marginal Analysis ([PMBA] Tsourapas & Frew, 2011), which assists decision-makers in identifying the most efficient use of resources, based on components such as need, efficiency, cost-effectiveness, and social validity.

The evidence from Chapters 3 and 4 and from previous research indicates that there are a number of barriers in the system that prevent educators accessing and using evidence to inform practice: lack of time, underdeveloped research literacy skills, and a lack of relevant, accessible and usable evidence (Cain, 2016; Gorard, 2020; Hemsley-Brown & Sharp, 2003; Van Schaik et al., 2018; Walker et al., 2019). To support schools to increase the use of evidence and uptake of evidence-based provision, researchers should provide schools with educator-friendly summaries of evidence that include evidence of cost-effectiveness and information to determine the compatibility of a programme and how it will function in their context. In addition, and aligned with the evidence, we propose that a more formalised and planned process should be developed in which trusted conduits, such as school improvement professionals, disseminate educator-friendly summaries of evidence to schools (Gorard et al., 2020; Nelson & Campbell, 2017; Nelson & O'Beirne, 2014; Slavin, 2019). To ensure that summaries of evidence are relevant and aligned with the priorities of schools, we recommend, as does Gorard et al. (2020), Nelson and O'Beirne (2014) and Schaik et al. (2018), that there needs to be more research collaborations between schools, school improvement organisations, and higher education institutions, to develop evidence that aligns with the needs and priorities of the schools.

Evidence from Chapters 5 suggests that typical in-service training alone does not lead to increased uptake of evidence-based methods. Durlak and DuPre (2008), Joyce and Showers (2002), Klingner et al. (2013) and Owen et al. (2021) have revealed that in-service training, that includes re-training and/or coaching is required to continue developing skills and self-efficacy. Although Owen et al. (2021) suggest that only a small amount of coaching is required. In line with this evidence, we propose that school improvement professionals work more closely with schools to identify opportunities to train educators in evidence-based methods during in-service training days, and to include follow-up training and coaching.

The evidence from Chapter 3, 4, 5, 6 along with existing research (Rogers, 2003), suggests that teachers are more likely to adopt evidence-based programmes perceived as compatible. The evidence from Chapter 6 indicates that adapting programmes to enhance compatibility is an effective way of increasing the use of evidence-based methods. The adaption of a programme is often viewed as a failure to implement a programme with fidelity (Durlak & DuPre, 2008). However, we argue that both fidelity and adaptions should be more frequently and rigorously evaluated through process evaluations and implementation research to identify the core components of a programme to retain to ensure effectiveness, and the elements that can be adapted without negatively impacting outcomes. We believe it is important that researchers work with schools to conduct these type of evaluations. In doing so, evidence-based programmes can be developed to adapt and fit broader contexts. In addition, and in line with the evidence building framework developed by Owen et al. (2022), we propose that greater efforts are required to more robustly field test programmes or teaching approaches with teaching staff during phase 1 (phase 1: define, manualise and test interventions), before moving to efficacy testing and later stages of implementation.

To develop a more robust and relevant evidence base, researchers should consider replicating the evaluation of school interventions (Chapter 3) in other regions across the UK to begin to create a database of the most commonly used programmes and approaches.

In addition, and in agreement with Gorard et al. (2020), our results indicate that more empirical research also needs to be funded and conducted to understand the best way to facilitate use of evidence in schools. Based on our findings, we would recommend that future research is conducted to determine optimum strategies to promote evidence-into-use, and the most effective in-service training and coaching models for schools.

Strengths

The data that makes up the empirical chapters of this thesis have contributed to the literature on determining how best to enhance the use of evidence-based methods in schools, which addresses the aims of GwE and Bangor University and wider Welsh Government objectives. In particular, Chapter 3 has provided important new information about the range of interventions in use in a cluster of school in Wales, which has already improved the quality of the provision in some of the cluster schools and enabled us to highlight areas that the cluster of schools might consider for future improvements to provision to help them towards a more evidence informed practice. The outputs of our research have also improved the quality of knowledge and decision-making by senior leaders.

Limitations

Throughout each empirical chapter we have discussed in detail the research limitations, therefore here we address the most pertinent limitations.

The review of evidence in Chapter 3 has some limitations that should be considered when interpreting and applying the results. First, although we had a clearly defined inclusion and exclusion criteria and conducted an explicit and rigorous systematic search of the literature, it cannot be assumed that all the available evidence was included. Therefore, the review cannot be considered exhaustive. Whilst the first author was aware of the negative effects of publication bias, this review chose to exclude the grey literature, because it was felt from a preliminary scan it would not improve the quality of the results. Also, because of resource constraints, it was not possible to employ a second reviewer to carry out an independent screening of the studies.

Whilst we were able to collect rich contextual data from a large cluster of schools in Wales, we recognise that data derived from a small number of participants in Chapters 4 and 5 may not necessarily generalise to other schools. However, most of our findings align very well with other research and theories, which increases the validity of our findings and the possibility that these findings might generalise to similar education settings. For the pilot evaluation of the adapted PATHS programme in Chapter 6, it was not possible to randomly assign participants to intervention and control groups and we cannot, therefore, assume that the adapted PATHS programme alone contributed to the improvement in learner outcomes. As a result of non-random assignment, there was a large imbalance in groups, and this may have affected results. However, the results from the pre-test data suggests that there was equivalence between groups on all measures at pre-test. To mitigate the effects from unequal groups, we calculated Hedge's g using the pooled standard deviation of the gain scores (Hedges, 2008). We acknowledge that future research using a more robust, experimental design that includes cluster randomisation at the level of the classroom, and cluster level analysis, is now required. We also recognise the need to collect other useful information such as pupil gender and free school meal information, as variable effects have been reported on these subset of pupils (Humphreys et al., 2018). Further limitation of this research design is the use of teacher-report measures, particularly as teachers were not blind to the conditions. It is also noteworthy that the measures of social and emotional competencies are not standardised, and this feature is likely to have inflated some of the effect sizes. Also, reported effect sizes are generally higher in experiments with younger populations (Wolf & Harbatkin, 2022).

Finally, in Chapters 5 and 6 we acknowledge there is a lack of long-term follow-up data to evaluate the potential adoption of the GBG and the PATHS programme in the following academic year. Evidence suggests that it can take multiple years for an innovation to embed within an organisation (Fixen et al., 2009; Rogers, 2003; Rohrbach et al., 2006). Unfortunately, as a result of the Coronavirus pandemic closing schools in March 2020, it was not possible to gather follow-up data. Consequently, we acknowledge that future research is required to evaluate the long-term impact of adoption and implementation of evidence-based methods.

Conclusion

The close-to-practice research outlined within this thesis was conducted with a cluster of schools in Wales. The schools had entered into a collaborative research partnership with regional school improvement consortia (GwE) and Bangor University to co-construct close-to-practice research that aligns with the values and priorities of the organisations involved and addresses the issues pertinent to the schools. The primary aim for the partnership was to increase the use of evidence-based methods in the cluster and move them towards a more evidence-informed practice. A priority outlined by the senior leaders in the schools was to improve their understanding of the evidence base for the interventions already in use, and our first research aim was to evaluate the range and evidence-base of interventions used across the cluster.

The main research aims of this thesis were to: explore ways to increase the use of evidence-based methods in schools (which included understanding what facilitates senior leaders decisions to adopt innovations); explore the effectiveness of educator-friendly reviews of evidence on existing provision; evaluate the extent that teachers implement evidence-based programmes following typical in-service training; and finally, to describe the process of contextually adapting an evidence-based programme to enhance compatibility and evaluate if this leads to increased adoption and higher levels of implementation.

The study conducted in Chapter 3 was the first to report empirical findings of the range and evidence base for interventions used in a cluster of schools, and we found that schools adopt a large number of interventions that have a very limited evidence base. The findings from Chapter 4 suggest that research evidence is infrequently used to inform the provision that schools adopt; in alignment with existing research, we found that school provision is largely informed by educators in other schools, the local school improvement advisor, and the perceived compatibility of programmes often facilitates adoption-decisions. We believe that the evidence from Chapter 3 and 4 indicate that the barriers to evidence-use persist– particularly the lack of time, and relevant, usable and accessible evidence. When exploring ways to increase the use of evidence-based methods we found that educator-friendly summaries of evidence have little impact on the provision that schools use (Chapter 3), and typical in-service training is not effective (Chapter 5). However, we found that the adaption of an evidence-based programme to make it more contextually compatible, does tend to lead to adoption and high levels of implementation (Chapter 6).

One of the key outcomes derived from Chapters 3, 5 and 6 is that the schools adopt too many programmes and strategies into use, which not only impacts school budgets, but also teachers' time and capacity to adopt and use other evidence-based methods and affects levels of implementation and intentions to sustain use. There was also strong evidence across Chapters 4, 5 and 6, which indicates that the perceived compatibility of programmes is an important facilitator of adoption decisions and the extent that a programme is implemented.

To facilitate the uptake of evidence-based methods into schools, we suggest that schools consider evaluating and refining the range of interventions in use to maximise their impact, and to allow for the adoption of other evidence-based methods that may be more cost-effective. We recommend greater use of collaborative research partnerships to develop evidence that is based on the needs and priorities of schools. We also suggest that educator-friendly summaries of evidence are created that include information that schools can use to determine the compatibility of a programme and how it will function in context. Lastly, researchers should work with schools to create adaptable evidence-based programmes to promote their compatibility with a broader range of contexts.

References

Ager, R., & Pyle, K. (2013). NFER Teacher voice omnibus: November 2012 Survey: Understanding Union Membership and Activity.

Appleton, P. (2000). Tier 2 CAMHS and its interface with primary care. *Advances in Psychiatric Treatment*, 6(5), 388-396.

Allen, G. (2011). *Early intervention: the next steps, an independent report to Her Majesty's government by Graham Allen MP*. The Stationery Office.

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., & Guyatt, G. H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of clinical epidemiology*, 64(4), 401-406.

Barrera, M., Berkel, C., & Castro, F. G. (2017). Directions for the advancement of culturally adapted preventive interventions: Local adaptations, engagement, and sustainability. *Prevention Science*, 18(6), 640-648.

Barrera Jr, M., & Castro, F. G. (2006). A heuristic framework for the cultural adaptation of interventions. *Clinical Psychology: Science and Practice*, 13(4), 311–316. https://doi.org/10.1111/j.1468-2850.2006.00043.x

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: Effects of individual contingencies for group consequences on disruptive behavior in a classroom 1. *Journal of applied behavior analysis*, 2(2), 119-124.

Bell, M., Cordingley, P., Isham, C., & Davis, R. (2010). *Report of professional practitioner use of research review: Practitioner engagement in and/or with research.* CUREE, GTCE, LSIS & NTRP.

Bell, B. S., & Kozlowski, S. W. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied psychology*, 93(2), 296.

British Educational Research Association. (2022). *Evidence-based practice in schools: What's really going on?* Retrieved https://www.bera.ac.uk/blog/evidence-based-practice-in-schools-whats-really-going-on

Bradshaw, J., & Main, G. (2016). *Child poverty and deprivation. In The well-being of children in the UK* (pp. 31-70). Policy Press.

Berry, V., Axford, N., Blower, S., Taylor, R. S., Edwards, R. T., Tobin, K., & 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.

Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational theory*, 57(1), 1-22.

Blank, C., & Shavit, Y. (2016). The association between student reports of classmates' disruptive behavior and student achievement. *Aera Open*, 2(3), 2332858416653921.

Blueprints for healthy human development (n.d.). https://www.blueprintsprograms.org/programs/20999999/good-behavior-game/print/

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, 36(5), 452-457.

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting positive behavior using the Good Behavior Game: A meta-analysis of single-case research. *Journal of Positive Behavior Interventions*, 18(3), 180-190.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, 12(3), 133-148.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. sage.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis?. *Qualitative research in psychology*, 18(3), 328-352. https://doi.org/10.1080/14780887.2020.1769238

British Educational Research Association (2018). *The BERA close-to-practice research report.* Retrieved online https://www.llse.org.uk/uploads/datahub/4567ceb%5E03in-03/2018-12-10-BERA-Close-to-Practice_report_Nov2018.pdf [accessed 15 October 2021]

Bristow, D., Carter, L., & Martin, S. (2015). Using evidence to improve policy and practice: the UK What Works Centres. *Contemporary social science*, *10*(2), 126-137.

Broekkamp, H. & Van Hout-Wolters, B. (2007) The gap between educational research and practice: a literature review, symposium and questionnaire, *Educational Research and Evaluation*,13, 203–220.

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions,* 12(3), 133-148.

Brown, C. (2017). Further exploring the rationality of evidence informed practice: A semiotic analysis of the perspectives of a school federation. *International Journal of Educational Research*, *82,* 28-39.

Brown, C., Daly, A. and Liou, Y-H. (2016) Improving Trust, Improving Schools: Findings from a Social Network Analysis of 43 Primary Schools in England, *Journal of Professional Capital & Community*, 1, 1, pp. 69-91.

Brown, C., & Greany, T. (2018). The evidence-informed school system in England: Where should school leaders be focusing their efforts?. *Leadership and Policy in Schools*, *17*(1), 115-137.

Brown, C., Schildkamp, K., & Hubers, M. D. (2017). Combining the best of two worlds: A conceptual proposal for evidence-informed school improvement. *Educational Research*, *59*(2), 154-172.

Brown, C., and Zhang, D. (2016) Is engaging in evidence-informed practice in education rational? Examining the divergence between teachers' attitudes towards evidence use and their actual instances of evidence use in schools, British Educational Research Journal, 42, 5, pp. 780-801.

Brown, C., & Zhang, D. (2017). Accounting for discrepancies in teachers' attitudes towards evidence use and actual instances of evidence use in schools. *Cambridge Journal of Education*, *47*(2), 277-295.

Bryman, A. (2006). Paradigm peace and the implications for quality. *International journal of social research methodology,* 9(2), 111-126.

Bryman, A., Becker, S., & Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research: A view from social policy. *International Journal of Social Research Methodology*, 11(4), 261-276.

Bumbarger, B., & Perkins, D. (2008). After randomised trials: issues related to dissemination of evidence-based interventions. *Journal of Children's Services*, 3(2), 55-64.

Cain, T. (2016). Research utilisation and the struggle for the teacher’s soul: A narrative review. *European Journal of Teacher Education*, 39(5), 616-629.

Campbell, M. J. (1980). The development of literacy in Bristol and Gloucestershire 1755-1870. University of Bath (United Kingdom).

Campbell, C., Pollock, K., Briscoe, P., Carr-Harris, S., & Tuters, S. (2017). Developing a knowledge network for applied education research to mobilise evidence in and for educational practice. *Educational Research*, *59*(2), 209-227.

Centre for the Study of Violence Prevention (2006) Model Programs Fact Sheet: Promoting Alternative Thinking Strategies. Boulder, CO: CSVP

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, 45(5), 283-292.

Ciarrochi, J., Atkins, P. W., Hayes, L. L., Sahdra, B. K., & Parker, P. (2016). Contextual positive psychology: Policy recommendations for implementing positive psychology into schools. *Frontiers in psychology*, 1561.

Coe, R. J. (2012). *Conducting your research. Research methods and methodologies in education*, 41-52.

Collaborative for Academic, Social, and Emotional Learning (2002). *Making the Case for Social and Emotional Learning and Service-Learning.* Chicago, IL: University of Illinois at Chicago.

Collaborative for Academic, Social and Emotional Learning (2003). Safe and sound: An educational leaders' guide to evidence-based social and emotional learning (SEL) programs. Chicago, IL: Author

Collaborative for Academic, Social and Emotional Learning (2015). CASEL schoolkit: A guide for implementing schoolwide academic, social, and emotional learning. Chicago, IL: Author

Coldwell, M., Greany, T., Higgins, S., Brown, C., Maxwell, B., Stiell, B., & Burns, H. (2017). *Evidence-informed teaching: an evaluation of progress in England. Research Report*. Department for Education.

Conduct Problem Prevention Research Group (CPPRG). (1995). *Psychometric properties of the social competence scale—teacher and parent rat*ings (Fast Track Project technical report). University Park, PA: Pennsylvania State University

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: a systematic review of randomised controlled trials in education research 1980–2016. Educational Research, 60(3), 276-291.

Coombes, L., Chan, G., Allen, D., & Foxcroft, D. R. (2016). Mixed-methods evaluation of the good behaviour game in English primary schools. *Journal of community & applied social psychology*, 26(5), 369-387.

Cooper, A., Levin, B., & Campbell, C. (2009). The growing (but still limited) importance of evidence in education policy and practice. *Journal of Educational Change*, 10(2-3), 159-171.

Corcoran, R. P., Cheung, A. C., Kim, E., & Xie, C. (2018). Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review*, 25, 56-72.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., & Coe, R. (2015). *Developing great teaching: lessons from the international reviews into effective professional development*. Project Report. Teacher Development Trust, London. https://dro.dur.ac.uk/15834/1/15834.pdf

Corrigan, A. (2002). *Social Competence Scale-Parent Version. Fast track project technical report.* Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.577.2894&rep=rep1&type=pdf

Craft, A. (1996). *Making School-Based INSET Effective*.

Cunningham, R. & Lewis, K., (2012). NFER *Teacher Voice Omnibus 2012 Survey: The use of the Pupil Premium*, Slough, NFER

CUREE (2010). *Understanding What Enables High Quality Professional Learning: a report on the research evidence*. CUREE.

Currie, J., & Thomas, D. (1999). Early test scores, socioeconomic status and future outcomes.

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

Davies, P. (1999) What is evidence-based education? *British Journal of Educational Studies*, 47, 2, pp. 108–121.

Denzin, N.K. (1978). *Sociological methods: A sourcebook*. New York, NY: McGraw-Hill.

Department for Education Excellence (1997) Excellence for all children: *Meeting Special Educational Needs*. London: HMSO.

Department for Education. (2010) *The importance of teaching*. London: Department for Education

Department for Education (2015) *Supporting the attainment of disadvantaged pupils: Briefing for school leaders*

Department for Education (2016a). *Educational Excellence Everywhere*. 9230, London:

Department for Education (2016b). *Educational Excellence Everywhere. Assessment of Impact*. DFE00086-2016, London: Department of Education.

Department for Education (2017) SEN support: research evidence on effective approaches and examples of current practice in good and outstanding schools and colleges,

Department for Education (2019) Special Educational Needs in England 2019

Dijkman, M. A., Harting, J., & van der Wal, M. F. (2015). Adoption of the Good Behaviour Game: an evidence-based intervention for the prevention of behaviour problems. *Health education journal*, 74(2), 168-182.

Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in school mental health promotion*, 1(3), 6-28.

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.

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.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of educational and psychological consultation*, 11(2), 193-221.

Domitrovich, C. E., Pas, E. T., Bradshaw, C. P., Becker, K. D., Keperling, J. P., Embry, D. D., & Ialongo, N. (2015). Individual and school organizational factors that influence implementation of the PAX good behavior game intervention. *Prevention Science*, 16(8), 1064-1074.

Donaldson, G. (2015). *Successful futures: Independent Review of Curriculum and Assessment Arrangements in Wales*. Retrieved from https://gov.wales/sites/default/files/publications/2018-03/successful-futures.pdf [Accessed 30 July 2021] Education Endowment Foundation (undated) Teaching and Learning Toolkit available at https://educationendowmentfoundation.org.uk/resources/teaching-learningtoolkit.

Dooley, K. E. (1999). Towards a holistic model for the diffusion of educational technologies: An integrative review of educational innovation studies. *Journal of Educational Technology & Society*, 2(4), 35-45. https://www.jstor.org/stable/jeductechsoci.2.4.35

Dusenbury, L., & Weissberg, R. P. (2017). Social emotional learning in elementary school: Preparation for success. *The Education Digest*, 83(1), 36.

Durlak, J. A. (2016). Programme implementation in social and emotional learning: basic issues and research findings. *Cambridge Journal of Education*, 46(3), 333-345.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology*, 41(3), 327-350.

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.

Education Endowment Foundation (2018) *EEF toolkit*. Available via EEF. https://educationendowmentfoundation.org.uk/news/eef-toolkit-new-survey-shows-59-of-senior-leaders-using-it-to-inform-their-/.

Education Endowment Foundation (2019) *'Putting Evidence to Work: A School's Guide to Implementation'*. Available via EEF. https://d2tic4wvo1iusb.cloudfront.net/eef-guidancereports/implementation/EEF_Implementation_Guidance_Report_2019.pdf?v=1635355218

Egan, D. (2007). Combating child poverty in Wales: are effective education strategies in place?. York: Joseph Rowntree Foundation.

Early Intervention Foundation (n.d.). The Good Behavior Game https://guidebook.eif.org.uk/programme/the-good-behaviour-game

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., ... & Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators.* Ascd.

Embry, D. D. (2002). The Good Behavior Game: A best practice candidate as a universal behavioral vaccine. *Clinical child and family psychology review*, 5(4), 273-297.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical child and family psychology review*, 11(3), 75-113.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications.* New York: Irvington Publishers.

Estyn (2020) Effective school support for disadvantaged and vulnerable pupils – case studies of good practice. Cardiff: Estyn. [Online]. Available from: https://www.estyn.gov.wales/system/files/2021-08/Effective%2520school%2520support%2520for%2520disadvantaged%2520and%2520vulnerable%2520pupils%2520en_0.pdf

Faria, A. M., Kendziora, K., Brown, L., O'Brien, B., & Osher, D. (2013). PATHS implementation and outcome study in the Cleveland Metropolitan School District. *American Institutes of Research*. Retrieved April, 30, 2020.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International journal of qualitative methods, 5(1), 80-92.

Ferrer-Wreder, L., Sundell, K., & Mansoory, S. (2012, April). Tinkering with perfection: Theory development in the intervention cultural adaptation field. *In Child & Youth Care Forum* (Vol. 41, No. 2, pp. 149-171). Springer US.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on social work practice*, 19(5), 531-540.

Fletcher, J. M., & Vaughn, S. (2009). Response to intervention: Preventing and remediating academic difficulties. *Child Development Perspectives*, 3(1), 30-37.

Fleuren, M., Wiefferink, K., & Paulussen, T. (2004). Determinants of innovation within health care organizations: literature review and Delphi study. *International journal for quality in health care*, 16(2), 107-123. https://doi.org/10.1093/intqhc/mzh030

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega Jr, R. (2014). Effects of the Good Behavior Game on challenging behaviors in school settings. *Review of educational research*, 84(4), 546-571.

Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: Developers’ views of implementation barriers and facilitators. *School Mental Health*, 1(1), 26.

Fuller, J. (2019) The confounding question of confounding causes in randomized trials, *British Journal for the Philosophy of Science*, 70(3), 901–926.

Furlong, J. (2015). Teaching tomorrow’s teachers. Retrieved from https://beta.gov.wales/sites/default/files/publications/2018-03/teaching-tomorrow’s-teachers.pdf [Accessed 29 July 2021]

Godfrey, D. (2014). Creating a research culture–lessons from other schools. *Seced*.

Godfrey, D. (2016). Leadership of schools as research-led organisations in the English educational environment: Cultivating a research-engaged school culture. Educational Management Administration & Leadership, 44(2), 301-321.

Godfrey, D. (2017) What is the proposed role of research evidence in England's 'self-improving' school system?. *Oxford Review of Education*, 43:4, 433-446, doi:10.1080/03054985.2017.1329718

Godfrey, D., & Handscomb, G. (2019). *Evidence use, research-engaged schools and the concept of an ecosystem. In An Ecosystem for Research-Engaged Schools* (pp. 4-21). Routledge.

Godfrey, D., & Brown, C. (2018). How effective is the research and development ecosystem for England's schools?. *London Review of Education*, *16*(1), 137-153.

Godfrey, D., Brown, C., & Stoll, L. (2017). *Leading for innovation and evidence-informed improvement.* Bloomsbury.

Goldacre, B. (2013) Building evidence into education. Accessed online http://dera.ioe.ac.uk/17530/7/ben%20goldacre%20paper_Redacted.pdf

Goldberg, J. M., Sklad, M., Elfrink, T. R., Schreurs, K. M., Bohlmeijer, E. T., & Clarke, A. M. (2019). Effectiveness of interventions adopting a whole school approach to enhancing social and emotional development: a meta-analysis. *European Journal of psychology of Education*, 34(4), 755-782.

Gorard, S. (2014). A proposal for judging the trustworthiness of research findings. Radical statistics, 110, 47-59.

Gorard, S. (2020). *Getting evidence into education: Evaluating the routes to policy and practice (Eds)*. Oxfordshire, UK: Routledge.

Gorard, S., See, B. H., & Siddiqui, N. (2014). *Switch-On Reading: Evaluation Report and Executive Summary*. Education Endowment Foundation.

Gorard, S., See, B. H., & Siddiqui, N. (2020). What is the evidence on the best way to get evidence into use in education?. *Review of Education*, *8*(2), 570-610.

Gottfredson, D. C., & Gottfredson, G. D. (2002). Quality of school-based prevention programs: Results from a national survey. *Journal of research in crime and delinquency*, 39(1), 3-35.

Gray, D. E. (2021). *Doing research in the real world*. Sage.

Greany, T. (2015). How can evidence inform teaching and decision making across 21,000 autonomous schools?. *Learning from the journey in England*. IoE Press.

Greany, T., & Brown, C. (2017). The Evidence Informed School System in England: Where should school leaders be focusing their efforts? *International Journal of Education Policy and Leadership*, 17, 1–23.

Greany, T., & Maxwell, B. (2017). Evidence-informed innovation in schools: aligning collaborative research and development with high quality professional learning for teachers. *International Journal of Innovation in Education*, *4*(2-3), 147-170.

Green, S., Higgins, J. P., Alderson, P., Clarke, M., Mulrow, C. D., & Oxman, A. D. (2008). Introduction. *Cochrane Handbook for Systematic Reviews of Interventions*.

Green, H, McGinnity, A, Meltzer, H, Ford, T, Goodman, R. Mental Health of Children and Young People in Great Britain, 2004. Report No.: 1403986371. *Office for National Statistics, 2005.*

Green, D. A., & Riddell, W. C. (2001). Literacy, numeracy and labour market outcomes in Canada (pp. 89-552). *Canada: Statistics Canada*.

Greenberg, M. T., Domitrovich, C., & Bumbarger, B. (2001). The prevention of mental disorders in school-aged children: Current state of the field. *Prevention & treatment*, 4(1), 1a.

Greenberg, M. T., Domitrovich, C. E., Graczyk, P. A., & Zins, J. E. (2005). The study of implementation in school-based preventive interventions: Theory, research, and practice. *Promotion of mental health and prevention of mental and behavioral disorders*, 3, 1-62.

Greenberg, M. T., Domitrovich, C. E., Weissberg, R. P., & Durlak, J. A. (2017). Social and emotional learning as a public health approach to education. *The future of children*, 13-32.

Greenberg, M. T., & Kusché, C. A. (1993). *Promoting social and emotional development in deaf children: The PATHS project*. University of Washington Press.

Greenberg, M. T., Kusche, C. A., Cook, E. T., & Quamma, J. P. (1995). Promoting emotional competence in school-aged children: The effects of the PATHS curriculum. *Development and psychopathology*, 7(1), 117-136.

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.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *The milbank quarterly*, 82(4), 581-629. https://doi.org/10.1111/j.0887-378X.2004.00325.x

Greenwood, C. R. (2001). Bridging the Gap between Research and Practice in Special Education: Issues and implications for Teacher Preparation. *Teacher Education and Special Education*, 24(4), 273–275. https://doi.org/10.1177/088840640102400402

Greenwood, C. R., Tapia, Y., Abbott, M., & Walton, C. (2003). A building-based case study of evidence-based literacy practices: implementation, reading behavior, and growth in reading fluency, K—4. *The Journal of Special Education*, 37(2), 95-110.

Groves, E. A., & Austin, J. L. (2017). An evaluation of interdependent and independent group contingencies during the good behavior game. *Journal of applied behavior analysis*, 50(3), 552-566.

Hallfors, D., & Godette, D. (2002). Will the ‘Principles of Effectiveness' improve prevention practice? Early findings from a diffusion study. *Health Education Research*, *17*(4), 461-470.

Hammersley, M. (2009). What is evidence for evidence-based practice? (pp. 101-111). *Brill Sense*.

Hammersley, M. (2015) Against ‘gold standards’ in research: On the problem of assessment criteria. *Was heißt hier eigentlich ‘Evidenz*. Available online at: www.degeval.de/fileadmin/users/Arbeitskreise

Hanushek, E. A., & Woessmann, L. (2020). Education, knowledge capital, and economic growth. *The economics of education*, 171-182.

Hargreaves, D. H. (2010). *Creating a self-improving school system.* Nottingham. National College for School Leadership

Hargreaves, A., & Fullan, M. G. (1992). *Understanding teacher development.* Teachers College Press, 1234 Amsterdam Avenue, New York, NY 10027.

Heath, B., Lakshmanan, A., Perlmutter, A., & Davis, L. (2010). Measuring the impact of professional development on science teaching: a review of survey, observation and interview protocols. *International Journal of Research & Method in Education*, 33(1), 3-20.

Hedges, L. V. (2008). What are effect sizes and why do we need them?. Child Development Perspectives, 2(3), 167-171.

Hedges, L. V., Pustejovsky, J. E., & Shadish, W. R. (2013). A standardized mean difference effect size for multiple baseline designs across individuals. *Research Synthesis Methods,* 4(4), 324-341.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288.

Hemsley-Brown, J., & Sharp, C. (2003). The use of research to improve professional practice: A systematic review of the literature. *Oxford Review of Education*, 29(4), 449-471.

Higgins, S, (2020). The development and worldwide impact of teaching and learning toolkit. In S. Higgins (Ed.), *Getting Evidence into Education: Evaluating the routes to Policy and Practice* (pp. 69 - 83). Oxfordshire: Routledge.

Higgins, S., Katsipataki, M., Villanueva-Aguilera, A. B., Coleman, R., Henderson, P., Major, L. E., & Mason, D. (2016). *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit.*

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288.

Hughes, C., & Cline, T. (2015). An evaluation of the preschool PATHS curriculum on the development of preschool children. *Educational Psychology in Practice*, 31(1), 73-85.

Humphrey, N., Barlow, A., & Lendrum, A. (2018). Quality matters: Implementation moderates student outcomes in the PATHS curriculum. *Prevention Science*, 19(2), 197-208.

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. *Journal of school psychology*, 58, 73-89.

Humphrey, N., Hennessey, A., Ashworth, E., Frearson, K., Black, L., Petersen, K., & Pampaka, M. (2018). Good Behaviour Game: Evaluation Report and Executive Summary. *Education Endowment Foundation*.

Inam, A., Tariq, P. N., & Zaman, S. (2015). Cultural adaptation of preschool PATHS (Promoting Alternative Thinking Strategies) curriculum for Pakistani children. *International Journal of Psychology*, 50(3), 232-239.

Institute for Effective Education (2019) *Engaging with evidence*. York: Institute for Effective Education

Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of educational research*, 79(1), 491-525.

Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American journal of public health*, 105(11), 2283-2290.

Jones, S. L., Procter, R., & Younie, S. (2015). Participatory knowledge mobilisation: an emerging model for international translational research in education. *Journal of Education for Teaching*, *41*(5), 555-573.

Joseph Rowntree Foundation (2020) Briefing: Poverty in Wales 2020. York [Online] Available at: https://www.jrf.org.uk/report/poverty-wales-2020

Joslyn, P. R., Donaldson, J. M., Austin, J. L., & Vollmer, T. R. (2019). The good behavior game: A brief review. *Journal of Applied Behavior Analysis*, 52(3), 811-815.

Joyce, B. R., & Showers, B. (2002). *Student achievement through staff development* (Vol. 3). Alexandria, VA: Association for Supervision and Curriculum Development.

Kam, C. M., Greenberg, M. T., & Kusché, C. A. (2004). Sustained effects of the PATHS curriculum on the social and psychological adjustment of children in special education. *Journal of emotional and behavioral disorders*, 12(2), 66-78.

Kam, C. M., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention science*, 4(1), 55-63.

Kam, C. M., Wong, L. W. L., & Fung, K. M. S. (2011). *Promoting social-emotional learning in Chinese schools: A feasibility study of PATHS implementation in Hong Kong*.

Khangura, S., Konnyu, K., Cushman, R., Grimshaw, J., & Moher, D. (2012). Evidence summaries: the evolution of a rapid review approach. *Systematic reviews*, 1(1), 1-9.

Kapoor, K. K., Dwivedi, Y. K., & Williams, M. D. (2014). Rogers' innovation adoption attributes: A systematic review and synthesis of existing research. *Information Systems Management*, 31(1), 74-91. https://doi.org/10.1080/10580530.2014.854103

Kellam, S. G., Mackenzie, A. C., Brown, C. H., Poduska, J. M., Wang, W., Petras, H., & Wilcox, H. C. (2011). The good behavior game and the future of prevention and treatment. *Addiction science & clinical practice*, 6(1), 73.

Kerssens-van Drongelen, I. (2001). The iterative theory-building process: rationale, principles and evaluation. *Management Decision*, 39(7), 503-512.

Khangura, S., Konnyu, K., Cushman, R., Grimshaw, J., & Moher, D. (2012). Evidence summaries: the evolution of a rapid review approach. *Systematic reviews,* 1(1), 1-9.

Klingner, J. K., Boardman, A. G., & McMaster, K. L. (2013). What does it take to scale up and sustain evidence-based practices?. *Exceptional Children*, 79(2), 195-211.

Kools, M., Pont, B., Gouedard, P., Rodriguez, T.., Stoll, L. & George, B. (2018). *Schools as learning organisations in Wales: A detailed analysis. In developing schools as learning organisations in Wales* (pp. 10–23). OECD

Publishing. https://doi.org/10.1787/9789264307193-7-en

Kools, M., Stoll, L., George, B., Steijn, B., Bekkers, V., & Gouëdard, P. (2020). The school as a learning organisation: The concept and its measurement. *European Journal of Education*, 55(1), 24-42.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241-253.

Kusché, C. A., & Greenberg, M. T. (2005). Promoting Alternative Thinking Strategies. https://content.pathsprogram.com/paths-current-users/502005-D-4c2c-DP-OT-EG.pdf

Lau, A. S. (2006). Making the case for selective and directed cultural adaptations of evidence-based treatments: Examples from parent training. *Clinical psychology: Science and practice*, 13(4), 295.

Levin, H. M. (2012). More than just test scores. *Prospects*, 42(3), 269-284.

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social science research*, 37(3), 888-902.

Machin, S., McIntosh, S., Vignoles, A., & Viitanen, T. (2001). *Basic skills, soft skills and labour market outcomes: secondary analysis of the National Child Development Study*. Department for Education and Employment Research Centre, Centre for the *Economics of Education, Research Report*, (250).

MacLeod, S., Sharp, C., Bernadinelli, D., Skipp, A. & Higgins S. (2015). Supporting the Attainment of Disadvantaged Pupils: Articulating Success and Good Practice. *Research Brief*.

Malin, J. R., Brown, C., Ion, G., van Ackeren, I., Bremm, N., Luzmore, R., & Rind, G. M. (2020). World-wide barriers and enablers to achieving evidence-informed practice in education: what can be learnt from Spain, England, the United States, and Germany?. *Humanities and Social Sciences Communications,* 7(1), 1-14.

Marchese, D. D., Becker, K. D., Keperling, J. P., Domitrovich, C. E., Reinke, W. M., Embry, D. D., & Ialongo, N. S. (2017). *A step-by-step guide for coaching classroom teachers in evidence-based interventions.* Oxford University Press.

Martens, B. K., Witt, J. C., Elliot, S. N., & Darveaux, D. X. (1985). *Professional Psychology: Research and Practice*.

McLafferty, S. L. (2003). Conducting questionnaire surveys. Key methods in geography, 1(2), 87-100.

McLeod, J. D., & Kaiser, K. (2004). Childhood emotional and behavioral problems and educational attainment. *American sociological review,* 69(5), 636-658.

Mitchell, E. S. (1986). Multiple triangulation: a methodology for nursing science. *Advances in nursing science*.

Mincu, M. E. (2015). Teacher quality and school improvement: what is the role of research?. *Oxford Review of Education*, 41(2), 253-269.

Movsisyan, A., Arnold, L., Copeland, L., Evans, R., Littlecott, H., Moore, G., & Rehfuess, E. (2021). Adapting evidence-informed population health interventions for new contexts: a scoping review of current practice. *Health research policy and systems*, 19(1), 1-19.

Movsisyan, A., Dennis, J., Rehfuess, E., Grant, S., & Montgomery, P. (2018). Rating the quality of a body of evidence on the effectiveness of health and social interventions: A systematic review and mapping of evidence domains. *Research Synthesis Methods,* 9(2), 224-242.

Murano, D., Sawyer, J. E., & Lipnevich, A. A. (2020). A meta-analytic review of preschool social and emotional learning interventions. *Review of Educational Research*, 90(2), 227-263.

NASUWT The Teachers' Union. (n.d.) Retrieved from https://www.nasuwt.org.uk/uploads/assets/uploaded/68ddda7f-afca-420c-a18798ea9271695f.pdf

National Teacher Research Panel (2010) *Habitats for teacher research: teacher perspectives on research as a sustainable environment for CPD*. Retrieved from http://www.curee.co.uk/files/publication/1313750504/NTRP%20survey%20report%20FINAL_0.pdf

Nelson, J., & Campbell, C. (2017). Evidence-informed practice in education: Meanings and applications. *Educational Research*, 59(2), 127–135. doi.org/10.1080/00131881.2017.1314115

Nelson, J. and C. O'Beirne. 2014. *Using Evidence in the Classroom: What Works and Why?* Slough: NFER [online]. Accessed January 23, 2017. https://www.nfer.ac.uk/publications/IMPA01

Nelson, J., Mehta, P., Sharples, J. and Davey, C. (2017). Measuring Teachers' Research Engagement: findings from a pilot study (EEF, London). Available online: https://educationendowmentfoundation.org.uk/public/files/Evaluation/Research_Use

Newby, P. (2014). *Research methods for education*. Routledge.

Ning, B., Van Damme, J., Van Den Noortgate, W., Yang, X., & Gielen, S. (2015). The influence of classroom disciplinary climate of schools on reading achievement: A cross-country comparative study. *School Effectiveness and School Improvement*, 26(4), 586-611.

Nolan, J. D., Houlihan, D., Wanzek, M., & Jenson, W. R. (2014). The Good Behavior Game: A classroom-behavior intervention effective across cultures. *School Psychology International*, 35(2), 191-205.

Nutley, S. M., Walter, I., & Davies, H. T. (2007). *Using evidence: How research can inform public services*. Policy press.

Nutley, S. M., Powell, A. E., & Davies, H. T. O. (2013). What counts as good evidence. Retrieved from https://researchrepository.standrews.ac.uk/bitstream/handle/10023/3518/What_Counts_as_Good_Evidence_published_version.pdf?sequence=1

OECD. (2010). Equity and Quality in Education: SUPPORTING DISADVANTAGED STUDENTS AND SCHOOLS. Retrieved from https://www.oecd.org/education/school/50293148.pdf

OECD. (2014). *Improving schools in Wales: An OECD perspective*. Retrieved from https://www.oecd.org/education/Improving-schools-in-Wales.pdf [accessed 12 October 2021]

OECD (2017). *The Welsh Education Reform Journey: A Rapid Policy Assessment*. Paris: OECD.

OECD (2018). *"Schools as learning organisations in Wales: A detailed analysis", in Developing Schools as Learning Organisations in Wales,* OECD Publishing, Paris. https://doi.org/10.1787/9789264307193-7-en

Ofsted. (2014). Below the radar: low-level disruption in the country's classrooms. https://www.gov.uk/government/publications/below-the-radar-low-level-disruption-in-the-countrys-classrooms

Owen, K. L., Hunter, S. H., Watkins, R. C., Payne, J. S., Bailey, T., Gray, C., ... & Hughes, J. C. (2021). Implementation Support Improves Outcomes of a Fluency-Based Mathematics Strategy: A Cluster-Randomized Controlled Trial. *Journal of Research on Educational Effectiveness*, 14(3), 523-542.

Owen, K. L., Watkins, R. C., & Hughes, J. C. (2022). From evidence-informed to evidence-based: An evidence building framework for education. *Review of Education*, 10(1), e3342.

Oxman, A. D., Lewin, S., Lavis, J. N., & Fretheim, A. (2009). SUPPORT Tools for evidence-informed health Policymaking (STP) 15: Engaging the public in evidence-informed policymaking. *Health research policy and systems*, 7(1), 1-9.

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.

Panayiotou, M., Humphrey, N., & Wigelsworth, M. (2019). An empirical basis for linking social and emotional learning to academic performance. *Contemporary Educational Psychology*, 56, 193-204.

Payton, J. W., Wardlaw, D. M., Graczyk, P. A., Bloodworth, M. R., Tompsett, C. J., & Weissberg, R. P. (2000). Social and emotional learning: A framework for promoting mental health and reducing risk behavior in children and youth. *Journal of school health*, 70(5), 179-185.

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). The Positive Impact of Social and Emotional Learning for Kindergarten to Eighth-Grade Students: Findings from Three Scientific Reviews.

Technical Report. *Collaborative for Academic, Social, and Emotional Learning* (NJ1).

Pegram, J., Watkins, R. C., Hoerger, M., & Hughes, J. C. (2022). Assessing the range and evidence-base of interventions in a cluster of schools. *Review of Education*, 10(1), e3336. https://doi.org/10.1002/rev3.3336

Ponce, O. A., & Pagán-Maldonado, N. (2015). Mixed methods research in education: Capturing the complexity of the profession. *International journal of educational excellence*, 1(1), 111-135.

Powers, J. D., Bowen, N. K., & Bowen, G. L. (2010). Evidence-based programs in school settings: Barriers and recent advances. *Journal of Evidence-Based Social Work*, 7(4), 313-331.

Pring, R. (2000). The 'false dualism' of educational research. *Journal of Philosophy of Education*, 34(2), 247-260.

Public Health England (2017) *Secondary school staff get mental health 'first aid' training*. https://www.gov.uk/government /news/secondary-school-staff-get-mental-health-first-aid-training

Ransford, C. R., Greenberg, M. T., Domitrovich, C. E., Small, M., & Jacobson, L. (2009). The role of teachers' psychological experiences and perceptions of curriculum supports on the implementation of a social and emotional learning curriculum. *School Psychology Review*, 38(4), 510.

Reid K. (2009). Improving attendance and behaviour in Wales: The action plan. *Educational Studies* 36(3): 233–247.

Rogers, E.M. (1995). *Diffusion of Innovations*. New York: Free Press.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation: transporting prevention interventions from research to real-world settings. *Evaluation & the health professions*, 29(3), 302-333.

Ross, S. M., & Morrison, J. R. (2021). Achieving better educational practices through research evidence: A critical analysis and case illustration of benefits and challenges. *ECNU Review of Education*, *4*(1), 108-127.

Ross, S. M., Sheard, M. K., Cheung, A., Elliott, L., & Slavin, R. (2011). Promoting primary pupils' social-emotional learning and pro-social behaviour: longitudinal evaluation of the Together 4 All Programme in Northern Ireland. *Effective Education*, 3(2), 61-81.

Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*. Sage.

Rycroft-Malone, J. (2004). The PARIHS framework—a framework for guiding the implementation of evidence-based practice. *Journal of Nursing Care Quality*, *19*(4), 297-304.

Substance Abuse and Mental Health Services Administration. (2011). National Registry of Evidence-Based Programs and Practice: PATHS, PATHS Preschool. Available at: http://www.nrepp.samhsa.gov/

Schaeffer, C. M., Bruns, E., Weist, M., Stephan, S. H., Goldstein, J., & Simpson, Y. (2005). Overcoming challenges to using evidence-based interventions in schools. *Journal of Youth and Adolescence*, 34(1), 15-22.

See, B.H. (2018) 'Evaluating the evidence in evidence-based policy and practice: examples from systematic reviews of literature, *Research in education*, 102 (1). pp. 37-61.

See, B. H., & Gorard, S. (2020). Effective classroom instructions for primary literacy: A critical review of the causal evidence. *International Journal of Educational Research*, 102, 101577.

Schaffer, K., Rouiller, S., Embry, D., & Ialongo, N. (2006). *The PAX good behavior game implementation rubric.* Unpublished technical report. Johns Hopkins University.

Sharples, J. (2013). Evidence for the Frontline. *London: The Alliance for Useful Evidence. Available at: www. alliance4usefulevidence. org/assets/EVIDENCE-FOR-THE-FRONTLINE-FINAL-5-June-2013. pdf.*

Shields, A., & Cicchetti, D. (1997). Emotion regulation among school-age children: the development and validation of a new criterion Q-sort scale. *Developmental psychology*, 33(6), 906.

Sklad, M., Diekstra, R., Ritter, M. D., Ben, J., & Gravesteijn, 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.

Sklad, M., & Park, E. (2017). Examining the potential role of education in the prevention of radicalization from the psychological perspective. Peace and Conflict: *Journal of Peace Psychology*, 23(4), 432.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15-21.

Slavin, R. E. (2008). Evidence-based reform in education: What will it take?. *European Educational Research Journal*, *7*(1), 124-128.

Slavin, R. E. (2017). Evidence-based reform in education. *Journal of Education for Students Placed at Risk*, *22*(3), 178–184. doi:10.1080/10824669.2017.1334560

Slavin, R. E. (2020) How evidence-based reform will transform research and practice in education, Educational Psychologist, 55:1, 21-31, DOI: 10.1080/00461520.2019.1611432

Sparks, S. (2018). *Teachers want education research. Education Week*, 27. Retrieved from https://www.edweek.org/leadership/teachers-want-education-research-the-feds-spend-millions-on-it-so-why-cant-it-get-to-the-classroom/2018/11?cmp=RSS-FEED

Stormont, M., Reinke, W., & Herman, K. (2011). Teachers' knowledge of evidence-based interventions and available school resources for children with emotional and behavioral problems. *Journal of Behavioral Education*, 20(2), 138-147.

Styles, B., & Torgerson, C. (2018) Randomised controlled trials (RCTs) in education research–methodological debates, questions, challenges, *Educational Research*, 60:3, 255-264, doi:10.1080/00131881.2018.1500194

Sutton Trust (2018). *School Funding and the Pupil Premium: NFER Teacher Voice Survey 2018.* https://www.suttontrust.com/wp-content/uploads/2019/11/TeacherPollingTables2019.pdf

Tarnowski, K. J., & Simonian, S. J. (1992). Assessing treatment acceptance: The abbreviated acceptability rating profile. *Journal of behavior therapy and experimental psychiatry*, 23(2), 101-106.

Tate, L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological rehabilitation*, *18*(4), 385-401.

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.

Thomas, G. (2012) Changing our landscape of inquiry for a new science of education, *Harvard Educational Review*, 82, 26–51.

Thomas, G. (2016). After the gold rush: Questioning the “gold standard” and reappraising the status of experiment and randomized controlled trials in education. *Harvard Educational Review*, *86*(3), 390-411.

Thomas, G. (2021). Experiment’s persistent failure in education inquiry, and why it keeps failing. *British Educational Research Journal*, *47*(3), 501-519.

Thurmond, V. A. (2001). The point of triangulation. Journal of nursing scholarship, 33(3), 253-258.

Tsourapas, A., & Frew, E. (2011). Evaluating ‘success’ in programme budgeting and marginal analysis: a literature review. *Journal of health services research & policy*, 16(3), 177-183.

Turner, A. J., Sutton, M., Harrison, M., Hennessey, A., & Humphrey, N. (2020). Cost-effectiveness of a school-based social and emotional learning intervention: evidence from a cluster-randomised controlled trial of the promoting alternative thinking strategies curriculum. *Applied health economics and health policy*, 18(2), 271-285.

Tyler, E., Watkins, R., Roberts, S., Hoerger, M., Hastings, R., Hulson-Jones, A., & Hughes, J. (2019). The collaborative institute for education research, evidence and impact: A case study in developing regional research capacity in Wales. *Wales Journal of Education*, 21(1), 89-108.

Vanderlinde, R., & van Braak, J. (2010). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. British educational research journal, 36(2), 299-316.

Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. Australian Journal of Advanced Nursing, The, 25(4), 40-48.

Van Lancker, W., & Parolin, Z. (2020). COVID-19, school closures, and child poverty: a social crisis in the making. *The Lancet Public Health*, 5(5), e243-e244.

Van Schaik, P., Volman, M., Admiraal, W., & Schenke, W. (2018). Barriers and conditions for teachers' utilisation of academic knowledge. *International Journal of Educational Research*, 90, 50-63.

Walker, H. M. (2004). Commentary: Use of evidence-based interventions in schools: Where we've been, where we are, and where we need to go. *School Psychology Review*, 33(3), 398.

Walker, M., Nelson, J., Bradshaw, S., & Brown, C. (2019). Teachers' engagement with research: what do we know? A research briefing. London: Education Endowment Foundation. Available at: https://educationendowmentfoundation. org. uk/public/files/Evaluation/Teachers_ engagement_with_research_Research_Brief_JK. pdf.

Weare, K. (2010). Mental health and social and emotional learning: Evidence, principles, tensions, balances. *Advances in school mental health promotion*, 3(1), 5-17.

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

Wehby, J. H., Maggin, D. M., Moore Partin, T. C., & Robertson, R. (2012). The impact of working alliance, social validity, and teacher burnout on implementation fidelity of the good behavior game. *School Mental Health*, 4(1), 22-33.

Welsh Government (2008), *School Effectiveness Framework: Building Effective Learning Communities Together*. Cardiff: Welsh Government

Welsh Government (2014a). Qualified for Life. Cardiff: Welsh Government.

Welsh Government (2015a) *Successful Futures: Independent Review of Curriculum and Assessment Arrangements in Wales.* Cardiff: Welsh Government.

Welsh Government (2015b) *Making best use of the Pupil Deprivation Grant funding*. Cardiff: Welsh Government. Retrieved from http://learning.gov.wales/resources/browse-all/pdg-what-really-works/?lang=en

Welsh Government. (2017). *Education in Wales: Our national mission. Action plan 2017-21.*

Welsh Government (2017) *Written statement—providing for the emotional and mental health needs of young people in school.* Retrieved from http://gov.wales/about/cabinet/cabinetstatements/2017/mentalhealthneeds/?lang=en

Welsh Government (2019). Retrieved from https://gov.wales/sites/default/files/statistics-and-research/2019-07/school-census-results-2019-764.pdf

Welsh Government (2021). *The National Strategy for Educational Research and Enquiry (NSERE): vision document.* Retrieved from https://gov.wales/sites/default/files/pdf-versions/2021/7/4/1626950790/national-strategy-educational-research-and-enquiry-nsere-vision-document.pdf

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.

Wiggins, G. (1998). *Educative Assessment. Designing Assessments To Inform and Improve Student Performance*. Jossey-Bass Publishers, 350 Sansome Street, San Francisco, CA 94104.

Williams, D., & Coles, L. (2007). Teachers' approaches to finding and using research evidence: An information literacy perspective. Educational research, 49(2), 185-206.

Willingham, D. T., & Daniel, D. (2021). *Making education research relevant.* Education Next. https://www. educationnext. org/making-education-research-relevant-how-researchers-can-give-teachers-more-choices/, last accessed, 6(06).

Wilson, R., Hemsley-Brown, J., Easton, C. & Sharp, C., 2003, Using Research for School Improvement: The LEA's Role, Slough, NFER Wilson, R., 2004, Taking control: how teachers use research, *NFER Topic*, Spring 2004/Issue 31

Wilson, R., 2004, Taking control: how teachers use research, *NFER Topic*, Spring 2004/Issue 31

Wisdom, J. P., Chor, K. H. B., Hoagwood, K. E., & Horwitz, S. M. (2014). Innovation adoption: a review of theories and constructs. *Administration and Policy in Mental Health and Mental Health Services Research*, 41(4), 480-502. https://doi.org/10.1007/s10488-013-0486-4

Wiskow, K. M., Matter, A. L., & Donaldson, J. M. (2019). The Good Behavior Game in preschool classrooms: An evaluation of feedback. *Journal of applied behavior analysis*, 52(1), 105-115.

Wolf, B., & Harbatkin, E. (2022). Making Sense of Effect Sizes: Systematic Differences in Intervention Effect Sizes by Outcome Measure Type. *Journal of Research on Educational Effectiveness*, 1-28.

Wrigley, T. (2018). The power of 'evidence': Reliable science or a set of blunt tools?. *British Educational Research Journal*, *44*(3), 359-376.

Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: Examining a positive variation of the good behavior game. *Journal of Positive Behavior Interventions*, 14(3), 173-180.

Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth: implications for prevention. *American Psychologist*, 67(4), 272.

Zins, J. E., & Elias, M. J. (2007). Social and emotional learning: Promoting the development of all students. *Journal of Educational and Psychological consultation*, 17(2-3), 233-255.

Table of Appendices

APPENDIX A: TABLE OF EXCLUDED INTERVENTIONS (CHAPTER 3)........1

APPENDIX B: WORKSHOP TRAINING EVALUATION FORM (CHAPTER 5)........3

APPENDIX C: EXAMPLE OF THE INTERVENTION RATING PROFILE – 15 (CHAPTER 5)........4

APPENDIX D: GOOD BEHAVIOR GAME – WEEKLY LOG (CHAPTER 5)........5

APPENDIX E: IMPLEMENTATION RUBRIC (CHAPTER 5)........6

APPENDIX F: EXAMPLE LESSON PLAN (CHAPTER 6)........7

APPENDIX G: EXAMPLE SCRIPT........8

APPENDIX H: EXAMPLE SUPPORTING MATERIALS........9

APPENDIX I: EXAMPLE TEACHER-REPORT IMPLEMENTATION CHECKLIST (CHAPTER 6)........10

APPENDIX J: EXAMPLE OBSERVATION CHECKLIST (CHAPTER 6)........11

Appendix A: Table of excluded interventions (Chapter 3)

Reason for excluding	Intervention
An accommodation or modification	Alternative playtime English Additional Language support Extended transition Individual Education Plan Managed Moves Personalised timetable Reduced hours Sensory breaks Visual Timetable Inclusion support
Provided by an external service	Barnardos Behaviour support services Behaviour support team CALDS (Child and Adolescent Learning Disability Services) CAMHS (Child and Adolescent Mental Health Services) Counselling services Educational Psychologist Family Engagement Officer Horse riding NWDAS drama workshop Occupational therapy practitioner School liaison officer activities TM Outreach TRAC services V IVA Wrexham FC Youth Offending Team YPB outreach
Staff or school facilities	Access staff Emotional/behavioural facilities First Contact Team Higher Level Teaching Assistant Key workers Learning coaches Non-teaching Assistant HOH Safeguarding Team
Pupils were not the recipients	Young carers in schools Behaviour diary
Universal/whole school provision/extra-curricular activity	5x60 sport activities After school clubs

Universal/whole school provision/extra-curricular activity	Pyramid club
	Good Behaviour game
	Art club
	Assembly
	Big Maths
	Boys lunch time group
	Canteen
	Challenge baby week
	Choir
	Creative Curriculum
	Daily Reading
	Dogs Trust - cannine welfare
	Don’t touch tell workshops
	Financial Education
	First aid
	Football Club
	Form period
	Form tutors
	Friday Skills Club
	Friendship groups
	G2G Lego
	Heads of house
	Homework Club
	Internet Safety
	Library
	Lunchtime Clubs
	Lunchtime games group
	Music lessons
	PHSE
	Pivotal
	Pop up provision
	Revision sessions
	Reward points/praise postcards
	School nurse
	School productions
	SEAL
	Sex and Relationships Education
	Subject staff
	Swimming
	Trips
	Visit to Danger Point
	Walkabout' Teacher
	Wellbeing afternoon activities
	Whole school attendance rewards
	Wide curriculum

Appendix B: Workshop training evaluation form (Chapter 5)

Workshop training evaluation form: Please select an answer to indicate what you thought about the training

Training session name: Good Behaviour Game Date: 22nd February 2018 ***Please circle as appropriate***

Category	Item	Rating				
		Strongly Disagree	**Disagree**	**Neither agree not disagree**	**Agree**	**Strongly Agree**
Overall	The training session was well organised	1	2	3	4	5
	The content of the session was covered in the time available	1	2	3	4	5
	Trainers provided all that I needed to complete training tasks	1	2	3	4	5
Process: materials	The presentation slides were relevant, clear and useful	1	2	3	4	5
	The handouts were relevant, clear and useful	1	2	3	4	5
Process: trainers	Trainer presentations were clear and concise	1	2	3	4	5
	Trainers demonstrated practical skills and knowledge	1	2	3	4	5
	Trainer feedback was clear and concise	1	2	3	4	5
Process: aspects	I learned from the presentation	1	2	3	4	5
	I learned from the modelling with group discussion	1	2	3	4	5
Outcomes	I know what ‘Good Behaviour Game’ aims to do	1	2	3	4	5
	I know how to use ‘Good Behaviour Game’ in the classroom	1	2	3	4	5
	I would like 1-1 training in the classroom in addition to today’s group session	Yes			No	

Any other comments:

………
………
………

Appendix C: Example of the Intervention Rating Profile – 15 (Chapter 5)

Primary Intervention Rating Scale: Teacher Version

Please read the following statements regarding implementation of the Good Behavior Game and circle the number that best describes your agreement or disagreement with each statement.

	Strongly Disagree	*Disagree*	*Slightly Disagree*	*Slightly Agree*	*Agree*	*Strongly Agree*
1. This was an acceptable intervention for the pupils' needs.	1	2	3	4	5	6
2. Most teachers would find this intervention appropriate for children with similar needs.	1	2	3	4	5	6
3. This intervention proved effective in supporting the pupils' needs	1	2	3	4	5	6
4. I would suggest the use of this intervention to other teachers.	1	2	3	4	5	6
5. The pupils' needs warrant use of this intervention.	1	2	3	4	5	6
6. Most teachers would find this intervention suitable for the needs of the pupils.	1	2	3	4	5	6
7. I used this intervention in the classroom setting.	1	2	3	4	5	6
8. This intervention did not result in negative side effects for the pupils.	1	2	3	4	5	6
9. This intervention was appropriate for a variety of pupils.	1	2	3	4	5	6
10. This intervention was consistent with those I have used in classroom settings.	1	2	3	4	5	6
11. The intervention was a fair way to handle the pupils’ needs.	1	2	3	4	5	6
12. This intervention was reasonable for the needs of the pupils.	1	2	3	4	5	6
13. I liked the procedures used in this intervention.	1	2	3	4	5	6
14. This intervention was a good way to handle pupils' needs.	1	2	3	4	5	6
15. Overall, this intervention was beneficial for pupils.	1	2	3	4	5	6

Appendix D: Good Behavior Game – Weekly log (Chapter 5)

Weekly GBG Log	Number of games played	Duration of each game
Week commencing 26/2/18		
Week commencing 5/3/18		
Week commencing 12/3/18		
Week commencing 19/3/18		
Week commencing 9/4/18		
Week commencing 16/4/18		
Week commencing 23/4/18		
Week commencing 30/4/18		
Week commencing 7/5/18		
Week commencing 14/5/18		
Week commencing 21/5/18		
Week commencing 4/6/18		
Week commencing 11/6/18		
Week commencing 18/6/18		
Week commencing 25/6/18		
Week commencing 2/7/18		
Week commencing 9/7/18		
Week commencing 16/7/18		

Appendix E: Implementation rubric (Chapter 5)

Good Behaviour Game Implementation Rubric

School name: ________________________________ Date ______________________

Teacher: ____________________________________ Researcher ____________________________

1. Pupils are in teams

Yes ☐ No ☐

2. Rules are displayed or pupil are reminded of rules

Yes ☐ No ☐

3. Rules are simple to understand and framed positively

Yes ☐ No ☐

4. Appropriate behaviours are monitored and tracked

Yes ☐ No ☐

5. Appropriate behaviour is verbally reinforced

Yes ☐ No ☐

6. The game is timed

Yes ☐ No ☐

7. Winner/s announced

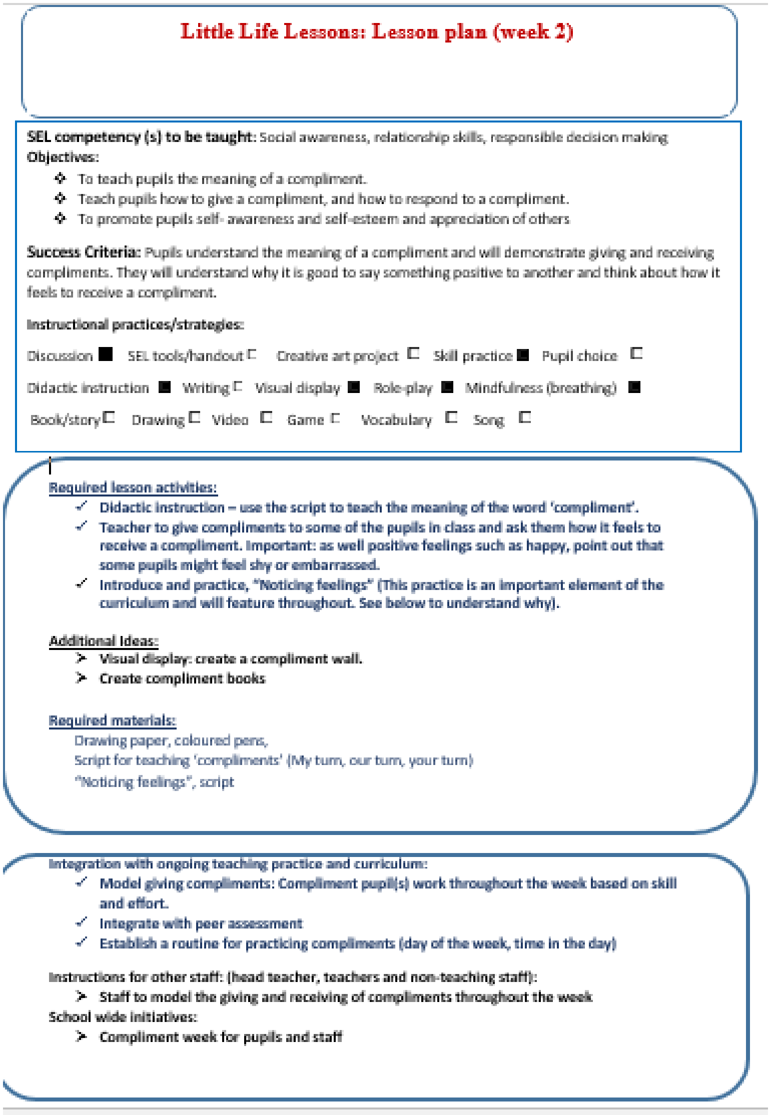
Yes ☐ No ☐

8. Rewards are delivered

Yes ☐ No ☐

Total points _________

Appendix F: Example lesson plan (Chapter 6)



Little Life Lessons: Lesson plan (week 2)

SEL competency (s) to be taught: Social awareness, relationship skills, responsible decision making

Objectives:

- To teach pupils the meaning of a compliment.
- Teach pupils how to give a compliment, and how to respond to a compliment.
- To promote pupils self- awareness and self-esteem and appreciation of others

Success Criteria: Pupils understand the meaning of a compliment and will demonstrate giving and receiving compliments. They will understand why it is good to say something positive to another and think about how it feels to receive a compliment.

Instructional practices/strategies:

Discussion ■ SEL tools/handout ☐ Creative art project ☐ Skill practice ■ Pupil choice ☐

Didactic instruction ■ Writing ☐ Visual display ■ Role-play ■ Mindfulness (breathing) ■

Book/story ☐ Drawing ☐ Video ☐ Game ☐ Vocabulary ☐ Song ☐

Required lesson activities:

- ✓ **Didactic instruction – use the script to teach the meaning of the word 'compliment'.**
- ✓ **Teacher to give compliments to some of the pupils in class and ask them how it feels to receive a compliment. Important: as well positive feelings such as happy, point out that some pupils might feel shy or embarrassed.**
- ✓ **Introduce and practice, "Noticing feelings" (This practice is an important element of the curriculum and will feature throughout. See below to understand why).**

Additional Ideas:

- ➢ **Visual display: create a compliment wall.**
- ➢ **Create compliment books**

Required materials:

Drawing paper, coloured pens,
Script for teaching 'compliments' (My turn, our turn, your turn)
"Noticing feelings", script

Integration with ongoing teaching practice and curriculum:

- ✓ **Model giving compliments: Compliment pupil(s) work throughout the week based on skill and effort.**
- ✓ **Integrate with peer assessment**
- ✓ **Establish a routine for practicing compliments (day of the week, time in the day)**

Instructions for other staff: (head teacher, teachers and non-teaching staff):

- ➢ **Staff to model the giving and receiving of compliments throughout the week**

School wide initiatives:

- ➢ **Compliment week for pupils and staff**

Appendix G: Example script

Teach the rules/expectations (script, week 1)

("My turn, our turn, your turn")

Teacher (My turn), step 1: Identify the expectation to teach and describe it

"The first rule is, to listen. You should listen when someone is talking. When you are listening you are looking at the person that is talking, you are quiet and not talking to anyone else. Listening is important in the classroom, dinner hall, corridor, and in assembly."

Teacher (My turn), step 2: List a rationale for teaching the behavior (Why is it important?)

"It is important to listen because it shows respect to others. It tells the teacher we are ready to learn. If we listen, we know what to do next. Listening helps us to stay safe. "

Teacher (My turn), step 3: Identify examples and non-examples (What would the behavior look/sound like? What would the behavior not look/sound like?)

Examples	Non-examples
• *Look at the person who is talking* • *Sitting still* • *Being quiet* • *Following instructions*	• *Looking at someone else* • *Fidgeting* • *Talking or giggling* • *Not following instructions*

Teacher (My turn), step 4: Practice/Role Play Activities

Model expected behavior (My turn): Teacher(s) models the behavior by reading a story/scenario. The teacher discusses why the story/scenario is an example of following the rule and why it's good to follow the rules and the consequences of not.

(Use the provided story, or select your own story).

Lead students through behavior (Our turn): Teacher and class role-play a classroom scenario of following the rule: teacher(s) and pupils will discuss as a class what the rule looks and sounds like. (Use the picture and role-play). Prompt the class to think about why it is good to follow the rule and the consequences for not following the rule.

Test to ensure students understand behavior (Your turn): Pupils will demonstrate the rule. Teacher will provide feedback and correct if necessary. If correction is required, repeat the process (my turn, our turn, your turn).

Step 5: Provide activities for practice and reteach throughout the week

Appendix H: Example supporting materials

Good sitting

Be a good friend.

A good friend helps.

Appendix I: Example teacher-report implementation checklist (Chapter 6)

Implementation checklist – lesson 1

Date of delivery? ___________ Approximate time taken to deliver a lesson? _________

Did you complete the required lesson activities?

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

To what extent did you follow the structure of required lesson activities outlined in the lesson plan?

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

Did you feel prepared for the lesson?

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

Did you integrate this week's lesson into your ongoing teaching practice?

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

Did 80-100% of pupils actively participate in the lesson activities?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

Did 80-100% of pupils display sustained interest and engagement during the lesson?

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐

Appendix J: Example observation checklist (Chapter 6)

Observation checklist

Date of observation: ______________ Teacher: __________________

Approximate time taken to deliver the lesson: _________

The lesson activities were completed

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

The teacher adhered to the structure of the lesson activities.

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

The teacher was prepared and enthusiastic when delivering the lesson.

Completely ☐ Mostly ☐ Somewhat ☐ A little ☐

80-100% of pupils displayed sustained interest and engagement during the lesson.

Completely ☐ Mostly ☐ Somewhat ☐ Not at all ☐