murdoch children's research institute

Social Ventures Australia
Funding • Investment • Advice

BAIN & COMPANY

THE RESTACKING ODDS

RESTACKING THE ODDS

TECHNICAL REPORT

The Early Years of School: An evidence-based restricted systematic review of school quality domains and intervention strategies

Carly Molloy
Ruth Beatson
Shona Cawley
Christopher Harrop
Sharon Goldfeld
Nicholas Perini

Project team

Restacking the Odds is a collaboration between three organisations, each with relevant and distinctive skills and resources:

Murdoch Children's Research Institute (MCRI) is an independent medical research institute. MCRI's research covers the breadth of health and medical research from basic science through to clinical sciences and population health. MCRI is committed to giving all children the opportunity to have a happy and fulfilled life.

- **Prof Sharon Goldfeld** – Director Centre for Community Child Health and Theme Director Population Health, Royal Children's Hospital and Murdoch Children's Research Institute.
- **Dr Carly Molloy** – Senior Research Officer and Senior Project Manager, Murdoch Children's Research Institute. Carly is the Project Manager for Restacking the Odds, and can be contacted at carly.molloy@mcri.edu.au

Bain & Company is one of the world's leading management consulting firms. Bain works with executives and organisations to help them make better decisions, convert those decisions into actions, and deliver sustainable success.

- **Chris Harrop** – Partner, and member of Bain's worldwide Board of Directors.

Social Ventures Australia (SVA) supports partners across sectors to increase their social impact. SVA helps business, government and philanthropists to be more effective funders and social purpose organisations to be more effective at delivering services

- **Nick Perini** – Director, SVA Consulting.

Our collective aspiration is to create a new approach to tackling intergenerational disadvantage in Australia that delivers ground-breaking results.

Suggested citation: Molloy C., Beatson R., Cawley, S., Harrop C., Goldfeld S., Perini N. Restacking the Odds – Technical Report: The Early Years of School: An evidence-based restricted systematic review of school quality domains and intervention strategies. Melbourne, Australia, 2020.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABRACADABA/ AbraA	A Balanced Reading Approach for Canadians Designed to Achieve Best Results for All (interactive web-based balanced reading program)
ATAR	Australian Tertiary Admissions Rank
A+PAAAC	Academic Achievement and Physical Activity Across the Curriculum intervention
BRIDGE	Bridging Mental Health and Education in Urban Schools
CAI	Computer assisted instruction
CBPR	Community-based participatory research
CDI	Child directed instruction
CIRC	Cooperative Integrated Reading and Composition
CPC	Child parent centres
CORI	Conceptual Oriented Reading Instruction
CRCT	Cluster randomised controlled trial
EYS	Early years of school (i.e. grade PK-3)
GBG	Good Behaviour Game
HPS	Health Promoting School
IYTCMP	Incredible Years Teacher Classroom Management Program
MA	Meta-analysis
MVPA	Moderate-vigorous physical activity
NGO	Non-government organisation
NR	Not reported
NOS	Not otherwise specified
PAAC	Physical Activity Across the Curriculum intervention
PACES	Partnerships for Active Children in Elementary School
PALS	Peer Assisted Learning Strategies
PATHS	Promoting Alternative Thinking Strategies
PCMP	Proactive Classroom Management Program
PD	Professional Development
PDA	Personal digital assistants
PE	Physical education
PISA	Programme for International Student Assessment
PK	Pre-kinder
QES	Quasi-experimental study
RCT	Randomised controlled trial
REE	Repeated exposure experiment
RQ	Research question
SEB	Socio-emotional and behavioural
SR	Systematic review
SWPBS	School-Wide Positive Behaviour Support
TCIT	Teacher child interaction training
TDI	Teacher directed instruction
TEAMS	Teachers Engaged in Authentic Mentoring Strategies
TR	Teacher reported
WHO	World Health Organisation
WIAT-II	Wechsler Individual Achievement Test 2nd Edition (standardised test for measuring academic achievement)
WITS	Walk away, Ignore, Talk, Seek help (a Community based, whole school peer victimisation prevention program that teaches children the following strategies for coping with bullying)

LIST OF STATISTICAL NOTATIONS

CI	Confidence interval
d	Cohens d
ES	Effect size
glass Δ	Glass Delta:
k	Number of studies
NR	Not reported
ns	Not significant
OR	Odds ratio
p	Significance level
r	Correlation coefficient
SD	Standard deviation
%	Percent
~	Approximately
+ve	Positive
-ve	Negative
β	Beta level
η^2	Eta-squared
<	Less than
>	Greater than

LIST OF RESULTS TABLE ABBREVIATIONS

C	Control
F/up	Follow up
I	Intervention
K	Kinder
min	minutes
Obs	Observed
r/ship	Relationship
Rel.	Relevant
Tx	Treatment
vs.	Versus

EXECUTIVE SUMMARY

Restacking the Odds: Project Background

Too many children are born into circumstances that do not provide them with a reasonable opportunity to make a good start in life. Disadvantaged circumstances for children lead to developmental inequities in physical health, social-emotional wellbeing, and academic learning. These inequities emerge in early childhood and often continue into adulthood, contributing to unequal rates of educational attainment, mental and physical health and income. In some cases, this experience is part of a persistent cycle of intergenerational disadvantage. Inequities constitute a significant and ongoing social problem and – along with the substantial economic costs – have major implications for public policy.

Research has shown that to redress these developmental inequities, effort delivered during early childhood (from pregnancy to 8 years of age) has the greatest benefit. As a result, *Restacking the Odds* focuses on five key evidence-based interventions/platforms in early childhood (*see Figure 1: Five Fundamental Strategies)*:

1. Antenatal care;
2. Sustained nurse home visiting;
3. Early childhood education and care;
4. Parenting programs; and
5. The early years of school.

These five strategies are only a subset of the possible interventions, but we have selected them carefully. They are notably *longitudinal* (across early childhood), *ecological* (targeting child and parent), *evidence-based, already available* in almost all communities, and able to be *targeted* to benefit the 'bottom 25 per cent'. Our premise is that by 'stacking' these fundamental interventions (i.e., ensuring they are all applied for a given individual) there will be a cumulative effect - amplifying the impact and sustaining the benefit.

For each of the five strategies, the intent is to use a combination of data-driven, evidence-based and expert-informed approaches to develop measurable, best practice indicators of quality, quantity (access) and participation (reach):

Quality: Are the strategies *delivered effectively*, relative to evidence-based performance standards? A high-quality strategy is one for which there is robust evidence showing it delivers the desired outcomes. A larger number of research studies have explored aspects of this question (i.e., "what works?") compared with quantity and participation. Therefore, we pay particular attention to the quality dimension in this report.

Quantity: Are the strategies *available locally* in sufficient quantity for the target population? "Quantity" helps us determine the quantum of effort and infrastructure needed to deliver the strategy adequately for a given population.

Participation: Do the *appropriately targeted* children and families *participate* at the right dosage levels? "Participation" shows us what portion of the relevant groups are exposed to the strategy at the level required to generate the desired benefit. (For example, attending the required number of antenatal

visits during pregnancy). Participation levels can be calculated whether the strategy is universal (for everyone), or targeted (intended to benefit a certain part of the population).

In this project, indicators of quality, quantity and participation are used to help identify gaps and priorities in Australian communities. This will include testing preliminary indicators in 10 communities over the next 3 years to determine which are pragmatic to collect, resonate with communities, and provide robust measures to stimulate community and government action.

The findings summarised in this report on the fifth strategic area – the Early Years of School - will provide essential inputs to guide subsequent work for the *Restacking the Odds* project. Because school is compulsory in Australian states and territories this report only covers the quality dimension. Participation is expected to be in line with state and territory legislation. There is a similar report for each of the five strategies.

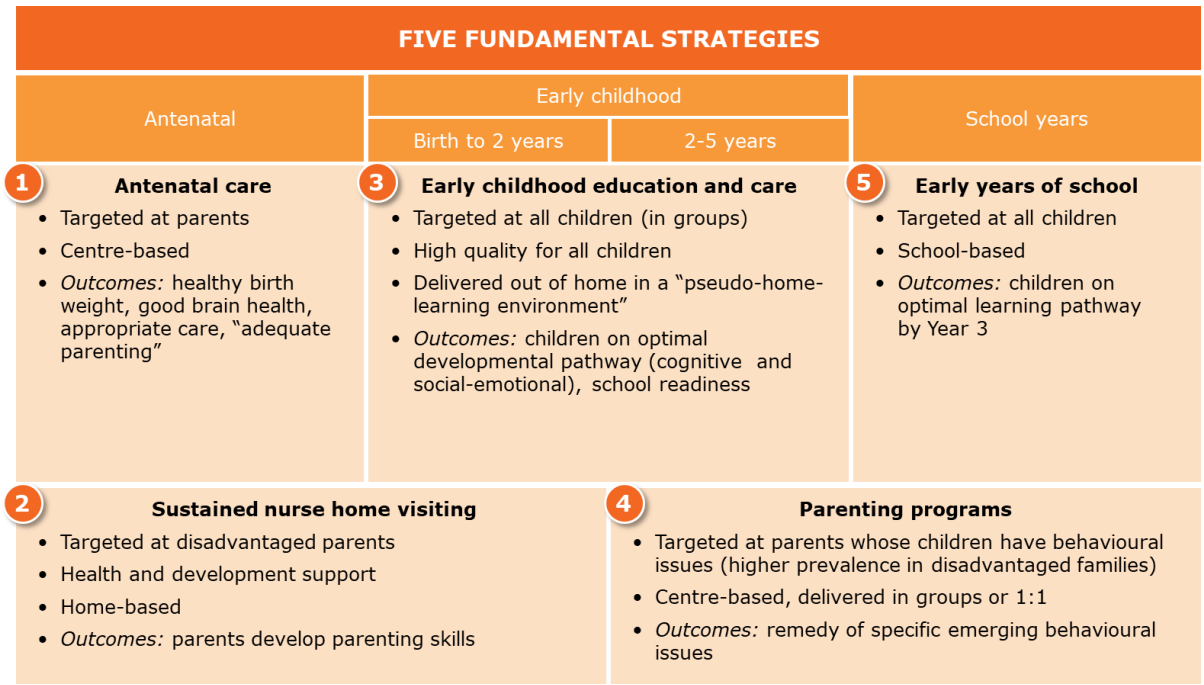

Figure 1: Five fundamental strategies

Introduction: The Early Years of School

Education is a fundamental social determinant of long-term health and quality of life (CSHD 2008; Cohen and Synne 2013; Hahn and Truman 2015). Strong and consistent evidence shows individual academic achievement is associated with a variety of health outcomes (e.g. risk and protective behaviours, morbidity, life-expectancy) and significantly influences access to employment and income (French et al. 2015; Hahn and Truman 2015; NSW CESE 2016). Education also has societal impacts, directly shaping the capabilities and productivity of future labour forces.

Internationally, educational attainment follows a social gradient. In both low and high-income economies, gaps in academic achievement typically show that children from disadvantaged backgrounds perform poorly relative to their socioeconomically advantaged counterparts (Carlisle and Murray 2015; Chung 2015; Sirin 2005; von Stumm 2017). Both the socioeconomic profile of individual students and schools is associated with academic performance in Australia (Lamb et al. 2015; Perry and McConney 2010). Critically, research shows the social gradient in cognitive development and academic

achievement is apparent early in childhood and widens with age (Chung 2015; Sirin 2005; von Stumm 2017). It therefore makes sense for school-based interventions to address these inequities early.

Universal education platforms are well-positioned to address socioeconomic inequities through increased access to higher quality schools (Carlisle and Murray 2015; Ladd and Loeb 2013). In Australia, each state and territory has its own framework for improving school quality. These frameworks describe a variety of approaches with differing improvement cycles that involve phases of evaluation, strategic goal setting, planning, implementation and monitoring. Each of the frameworks identify a range of domains thought to reflect school quality, and within each of the domains, may suggest a variety of improvement strategies. However, there are several concerns with the school quality evaluation tools utilised in existing frameworks. Limitations include overly complex structures, reliance on subjective ratings from school leaders, and ambiguity of quality indicators compromising the extent to which they are measurable and modifiable. Critically, the strength of evidence underpinning each of the quality domains identified has not been well documented. In some cases "what works" is not well known and in others, the strength of supporting evidence is not always clear. This situation does not leave the educator or school leadership with much direction on how to evaluate the relative importance or impact of an indicator or where they should focus their efforts.

If schools are to deliver high quality education with the aim of achieving equity in life outcomes related to academic achievement and health and well-being, it is important to know which strategies have demonstrated positive effects on child academic/cognitive, social-emotional, and health outcomes. An understanding of the strategies that significantly improve outcomes among children in the earliest grades is also critically important, given evidence that inequalities observed in the early years predict longer-term outcomes and prior research has demonstrated positive effects of early education interventions.

High quality education in *the Early Years of School* (EYS) is one of the five effective early intervention strategies identified by Restacking the Odds and thus is the focus of this review.

Aim

This review seeks to (a) identify effective school-based strategies to improve child outcomes (including academic achievement, social, emotional, and behavioural development), and (b) evaluate the evidence base specific to children in the early years of school, so that (c) a school quality assessment aligned with the evidence base can be developed, (d) educational decision-makers have the necessary information, and (e) gaps in the literature are identified to guide the direction of future research.

Method

We undertook a restricted systematic review; a research methodology that uses similar methods and principles to a comprehensive systematic review but makes concessions to the breadth and depth of the process, in order to be completed within a shorter timeframe. Rigorous methods for locating, appraising and synthesising the evidence related to a specific topic are utilised; however, the methodology places several limitations in the search criteria and in how the evidence is assessed. As formal schooling is compulsory in all Australian States and Territories from approximately 5 years of age, the search for the key drivers was restricted to those concerning quality.

Peer-reviewed literature

We sought to identify meta-analyses and systematic reviews of school-based interventions to improve student outcomes. Where no such publications were identified or those that were identified yielded low levels of evidence (e.g. correlational syntheses), the search was extended to randomised controlled trials (RCTs). In cases where no experimental studies met inclusion criteria, quasi-experimental studies

were also considered. A dual-phase approach to the search was applied. In the first phase search terms were kept broad and had the form: (synthesis terms) AND (school terms) AND (early childhood terms). In the second phase, terms specifically targeted 13 quality domains identified as common to existing school quality frameworks/evaluation tools. Restrictions were applied to the date of publication such that the Phase 1 search covered 2012 to 2018, whereas Phase 2 searches for research syntheses had no date restriction. In cases where it was necessary to search for RCT or quasi-experimental level evidence, the date of publication was restricted to 2008-2018. In contrast, where an extensive body of meta-analytic literature was identified for a specific domain, additional recency and relevance criteria were applied. Specifically, in two instances (domains 1 and 5) papers were selected only if published post 2010 and pooled effect sizes specific to the early years of school were presented.

Meta-analyses and systematic reviews constitute the highest levels of evidence, based on the NHMRC evidence hierarchy, as they combine results from multiple studies to increase statistical power (to detect effects) and produce a more precise estimate of the effect of treatment by consolidating sometimes conflicting results across studies (Hoffman 2015). RCTs on the other hand are considered the 'gold standard' way to assess a program's effectiveness.

Ranking the evidence

Each systematic review, meta-analysis, and RCT or quasi-experimental study meeting inclusion criteria was subject to a quality and bias check. Study quality includes assessment of internal validity (or the degree to which the design and conduct of the study avoid bias) and external validity (or the extent to which the results of the study can be applied, or generalised, to the population outside the study). The quality and bias information was used to consider the conclusions of included studies and the potential effectiveness of each strategy identified within each quality domain.

Considering the accumulated evidence, a judgement was reached about the strength of the evidence base for each quality domain (See Appendix F). The criteria was adapted from The California Evidence-based Clearinghouse for Child Welfare (CEBC 2016). Strength of evidence assessments were made by two independent raters and consensus was reached in cases where there was a discrepancy in initial ratings.

- *Well Supported*. Clear and consistent evidence of benefit across multiple strategies.
- *Supported.* Clear evidence of benefit for at least one strategy.
- *Promising*. Evidence suggestive of benefit for multiple strategies but more evidence needed.
- *Preliminary.* Evidence suggestive of benefit for at least one strategy but more evidence needed.
- *Mixed.* Conflicting findings for similar strategies.
- *Unknown*. Insufficient evidence to determine whether identified strategies are beneficial.
- *Not Supported*. Evidence consistently demonstrates identified strategies are ineffective or concerning.

Expert evaluation of draft indicators

The list of indicators was vetted by *Dr John Ainley,* Principal Research Fellow, Australian Council for Educational Research. Dr Ainley was asked to independently comment on (a) the list of Supported and Well Supported EYS quality domains and (b) the corresponding quality indicators.

Findings

The literature search and screening process resulted in the identification of 83 relevant publications, of which 66 were meta-analyses or reviews, providing the highest-levels of evidence. An evaluation of the evidence-base was conducted for each of the 13 domains, identifying 5 Well Supported and 4 Supported quality domains. Within these domains, the review identified 21 general strategies that have demonstrated effectiveness for children in the early years of school. These findings informed the development of evidence-based indicators to establish school quality.

Quality indicators

In total, 37 quality indicators were developed. These indicators are tied to school processes and teaching staff competencies that map to Well Supported and Supported quality domains. For the Well Supported domains, indicators relate to application of pedagogical content knowledge, differentiated teaching practices, use of technology, social-emotional development supports, and optimal professional development. Quality indicators for Supported domains relate to peer teaching, the incorporation of physical activity in the school day, class size, and partnerships with families.

Conclusion

Overall, the review indicates that there is a reasonably strong evidence base supporting several of the domains identified in existing school quality frameworks (nine of thirteen identified domains were rated Supported or Well Supported). The review also shows many of the strategies underpinning these domains have demonstrated effectiveness for children in the early years of school. The identification of these strategies together with the strength of evidence assessment for each provides a useful resource for guiding school selection of quality improvement initiatives.

Implications

The developed indicators will help identify gaps and priorities for Australian schools. We will test them in several Australian communities to determine which are pragmatic to collect, resonate with school communities, and provide robust measures to stimulate community and government action. We have followed a similar path for the other four fundamental strategies that Restacking the Odds is focusing on – antenatal care, sustained nurse home visiting, parenting programs, and early childhood education and care.

BACKGROUND: RESTACKING THE ODDS

Too many children are born into circumstances that do not provide them with a reasonable opportunity to make a good start in life. Disadvantaged circumstances for children lead to developmental inequities in physical health, social-emotional wellbeing, and academic learning – that is, differential outcomes that are preventable.

Inequities emerging in early childhood often continue into adulthood, contributing to unequal rates of low educational attainment, poor mental and physical health and low income. In some cases, this experience is part of a persistent cycle of intergenerational disadvantage. Inequities constitute a significant and ongoing social problem and – along with the substantial economic costs – have major implications for public policy.

The importance of early childhood and the impact of this period on long-term developmental outcomes has been well documented (Moore et al. 2017). Research has demonstrated that this period is crucial for brain development across all domains, and that both risk and protective factors encountered by the child during this time can have life-long impacts (Walker et al. 2011). In particular, exposure to multiple risk factors predicts more severe, adverse developmental consequences compared with a singular risk factor (e.g. Ferraro and Shippee 2009; Trentacosta et al. 2008). Furthermore, research has shown that developmental interventions that isolate only one risk factor are less likely to work than those that are multi-faceted (e.g. James et al. 2016; Nigg, Allegrante, and Ory 2002; Nigg and Long 2012).

The premise behind the Restacking the Odds approach to intervention is that resources/assets accumulate and the benefits of multiple assets accrue, leading to more positive outcomes. In line with this premise and research on cumulative risk, we hypothesise that inequities can be reduced by using existing, evidence-based interventions and approaches from service providers of the following five strategies: antenatal care; sustained nurse home visiting; early childhood education and care; parenting programs; and the early years of school. These strategies are notably longitudinal (across early childhood), ecological (targeting child and parent), evidence-based (RCT level support), and able to be targeted (aimed at benefiting the 'bottom 25 per cent', namely the most disadvantaged). By 'stacking' these fundamental interventions (i.e., ensuring they are all applied) it is predicted that there will be a cumulative effect, amplifying the effect and resulting in sustained benefits.

In order to achieve this, the *Restacking the Odds* project seeks to use the existing evidence within the 5 fundamental strategies of early childhood, to develop best practice benchmark frameworks that better define indicators of quality, access (quantity), and reach (participation).

This report focuses on the strategy of increasing quality education in *the Early Years of School*. There is a similar report for each of the five strategies.

INTRODUCTION: THE EARLY YEARS OF SCHOOL

Education is a fundamental social determinant of long-term health and quality of life (CSHD 2008; Cohen and Synne 2013; Hahn and Truman 2015). It is recognised as both an element of- and cause of-health status (Hahn and Truman 2015). Strong and consistent evidence shows individual academic achievement is associated with a variety of health outcomes (e.g. risk and protective behaviours, morbidity, life-expectancy) and significantly influences access to employment and income (French et al. 2015; Hahn and Truman 2015; NSW CESE 2016). Education also has societal impacts, directly shaping the capabilities and productivity of future labour forces. Indeed, educational attainment has been linked to national economic performance (OECD 2013) and government spending (Levin et al. 2007), social capital outcomes (e.g. level of trust in others, political efficacy, community involvement) (Rhodes, Cordie, and Wooten 2019) and level of participation in criminal activity (Lochner and Moretti 2004; Levin et al. 2007).

Internationally, educational attainment is distributed across a social gradient. In both low and high-income economies, gaps in academic achievement typically show that children from disadvantaged backgrounds perform poorly relative to their socioeconomically advantaged counterparts (Carlisle and Murray 2015; Chung 2015; Sirin 2005; von Stumm 2017). Lower academic achievement can reduce access to higher education and employment opportunities, leading to future difficulty in accessing health and economic resources (Hahn and Truman 2015; Marmot 2018). Indeed, in Australia, there is evidence that students from low socioeconomic backgrounds are less likely to complete Year 12 and/or receive an ATAR, and when they do, average ATAR scores are much lower, impacting access to higher education (Lamb et al. 2015; ACARA 2016). The socioeconomic profile of schools (in addition to the SES of individual students) is also associated with academic performance in Australia (Perry and McConney 2010).

Critically, research shows the social gradient in cognitive development and academic achievement is apparent early in childhood (Chung 2015; Sirin 2005; von Stumm 2017). Australian data shows substantial academic achievement gaps. Recent Australian Curriculum, Assessment, and Reporting Authority (ACARA 2018) reports show that grade 3 NAPLAN achievement in reading, writing, spelling, grammar and punctuation, and mathematics are negatively related to parent education and occupation status, across all states and territories[1]. For example, 98% of students whose parents have a Bachelor-degree or higher perform at or above national minimum standards in reading compared with 86% of children whose parents' highest qualification is year 11. Similarly, for students whose parents are not in paid work, 84-87% achieved at or above national minimum standards, compared with 96-99% of children whose parents were in the two highest occupational groups, across all academic domains.

Fortunately, there is evidence that some of the negative effects of low SES in childhood can be buffered by increasing education quality (Barnett 2011; Carlisle and Murray 2015). Early randomised controlled trials investigating the effectiveness of increasing education quality for socioeconomically disadvantaged preschool and kindergarten children (e.g. the High/Scope Perry Preschool Program and Carolina Abecedarian Project of the 1960s and 1970s) showed long-term benefits for both educational attainment and health outcomes (Muennig et al. 2009; Ramey et al. 2000). In Australia, research shows that the academic quality[2] of schools interacts with student socioeconomic status (Lim, Gemici, and Karmel 2014). Using 2006 cohort data from the Australian Longitudinal Studies of Australian Youth, Lim et al (2014) found that a significant gap (in PISA scores) between students from low and high SES

[1] In early cycles of the NAPLAN there were large amounts of missing data for parental education and occupation, but coverage in recent cycles has reached acceptable levels.

[2] School academic quality was defined by modelling the predicted TER scores and probability of an 'average' students attending university by 19 years of age using several characteristics of each school

backgrounds occurs in low-quality secondary schools, but that this gap disappears at high-quality schools. This research indicated that the impact of attending a high-quality school relative to low-quality schools more than doubles the chance of completing year 12 among low performing low SES students.

Given that differences in the academic performance of children from different SES backgrounds emerge early and widen with age (Chung 2015; Sirin 2005; von Stumm 2017), it makes sense that school-based interventions to address these inequities also occur early. In Australia, attendance at school is compulsory at age 5 or 6 dependent on individual state and territory mandated by law (Krieg and Whitehead 2015; ANZHES n.d.). The early years of school, defined as the Foundation[3] Year through to Year 3 (Hard and O'Gorman 2007; Jay, Knaus, and Hesterman 2014) is a critical time to develop children's language, cognition, social-emotional functioning, and generally prepare them for the acquisition of knowledge and skills required for a successful life (Bennett and Tayler 2006). In the formal school setting there is an explicit emphasis on further developing children's language, literacy and numeracy skills (Harrison et al. 2010; Laevers 2005). Yet to give all children the best life prospects it is important that they receive good quality education. Indeed, evidence shows that early school-based intervention programs that target all students and high-risk subgroups such as those from disadvantaged families have great potential to reduce inequities in child development (Smith et al. 2016; Sklad, Diekstra, Ritter, et al. 2012; Piasta and Wagner 2010; Barnett 2011; Dietrichson et al. 2017)

The available research suggests universal education platforms are well-positioned to address socioeconomic inequities through increased access to higher quality schools (Carlisle and Murray 2015; Ladd and Loeb 2013). While, there is a vast literature that examines specific school-quality factors, such as small class size, teacher-training, family-partnerships, and child outcomes there are significant challenges to the conceptualisation and measurement of school quality (Ladd and Loeb 2013) (Ladd and Loeb 2013). To this end there is a paucity of research that has tested/developed innovative ways to measure of school quality (Cohen and Synne 2013).

Australia has a National Quality Standard (NQS) for the early childhood education and care sector but not for school. The NQS was developed and implemented across the ECEC sector as a way to improve its quality. There are 15 quality standards that cover seven quality areas: educational program and practice, children's health and safety, physical environment, staffing arrangements, relationships with children, collaborative partnerships with families and communities, governance and leadership. Some educational authorities and researchers have argued that the NQS is also appropriate for use in the early years of school settings (Bope and Barblett 2016; Australian Children's Education & Care Quality Authority 2009). In practice, the NQS has been implemented in all primary schools from preschool to Year 2 since 2016 in Western Australia (Bope and Barblett 2016).

Given the importance of the early years of school to children's health and quality of life it is somewhat surprising that Australia does not have an accredited national quality framework to guide school quality performance. There are state-based approaches and other tools available such as the Victorian Framework for Improving Student Outcomes (Department of Education and Training 2019) and New South Wales School Excellence Framework (NSW Department of Education 2019) that differ in terms of structure, rating assessment and detail. However, there are concerns with the applicability of some frameworks. Limitations include overly complex structures, reliance on subjective ratings from school leaders, and ambiguity of quality indicators compromising the extent to which they are measurable and modifiable. Although extant frameworks identify a range of indicators considered important to educational quality, in some cases "what works" is not well known and in others, the strength of

[3] The Foundation Year is the first year of formal schooling, also referred to as the Preparatory year or Pre-Primary in some Australian states, Reception in the United Kingdom, and Kindergarten in the United States.

supporting evidence is not always clear. Traditional approaches to measuring school quality (such as measures of resources, direct observation of classroom interactions, and test performance) have a variety of challenges and limitations (Ladd and Loeb 2013). For example, measures of resources do not capture how effectively resources are deployed, direct observations of classroom interactions are resource intensive and evaluator ratings are difficult to standardise, and measures of test performance often limit conceptualisation of quality to specific subject areas. Measures of test performance as proxies for school quality are also problematic in that they often fail to disentangle the contribution of the school from the contribution of the student background, or indeed student entry capabilities. This situation does not leave the educator or school leadership with much direction on how to evaluate the relative importance or impact of an indicator or where they should focus their effort. Evidence-based quality indicators that are practical to measure are required to help educators and school leaders focus their efforts. Quality frameworks should be informed by high levels of evidence and focus on modifiable factors that have the greatest impact on child development.

Educational decision-makers (from teachers through to government) invest significant resources in the provision of education and student well-being. If schools are to deliver high quality education with the aim of achieving equity in life outcomes related to academic achievement and health and well-being, it is important to know which strategies have demonstrated positive effects on child academic/cognitive, social-emotional, and health outcomes. An understanding of the strategies that significantly improve outcomes among children in the earliest grades (Foundation-3) is also critically important, given evidence that inequalities observed in the early years predict longer-term outcomes and prior research has demonstrated positive effects of early education interventions (e.g. Wanzek et al. 2016). Such knowledge can be used to guide the development of (a) school policy and professional development strategies with the best chance of achieving the desired outcomes and (b) measures to ensure continuous quality improvement in an Australian service system context.

In the interest of facilitating easy translation of our findings into regular practice in Australian schools, we considered structuring our search for evidence explicitly around either the FISO or the NSIT. However, since the FISO is a strategy of the Victorian Department of Education and Training, framing the evidence strictly within that context may have limited the relevance and applicability of our findings to other Australian States and Territories. On the other hand, although the NSIT is available to schools nationally, our consultations with experts indicated that uptake has not been widespread to date. Therefore, rather than basing our search on a single framework, we constructed a list of school quality domains that was informed by all of the frameworks. As such, the list of school quality domains is related to each of the frameworks we identified, without being strictly bound to any specific framework.

AIM

High quality education *in the Early Years of School* (EYS) is one of the five effective early intervention strategies identified by Restacking the Odds and thus is the focus of this review.

The aims were to (a) identify effective school-based strategies to improve child outcomes (including academic achievement, social, emotional, and behavioural development), and (b) evaluate the evidence base specific to children in the early years of school, so that (c) a school quality assessment aligned with the evidence base can be developed, (d) educational decision-makers have the necessary information, and (e) gaps in the literature are identified to guide the direction of future research.

Consistent with RSTO reviews of other strategies (i.e. antenatal care; sustained nurse home visiting; early childhood education and care; and parenting programs) this review emphasises questions related to quality. It seeks to identify which school-based strategies are significantly related to better child

outcomes, and correspondingly, which process indicators can be used to measure the provision of high-quality education. Because school is compulsory, and delivered universally, it does not investigate questions of quantity (e.g. in what quantity should quality EYS be available?) or participation (e.g. what population is most likely to benefit from participation in EYS?).

METHOD:

This literature review utilised a targeted restricted evidence assessment (REA) methodology. The REA is a research methodology that uses similar methods and principles to a systematic review but makes concessions to the breadth and depth of the process, in order to be completed within a short timeframe. Rigorous methods for locating, appraising and synthesising the evidence related to a specific topic are utilised; however, the methodology places a number of limitations in the search criteria and in how the evidence is assessed.

The first step in conducting this review involved scoping domains considered relevant to quality early education in the early years of school, with particular emphasis on current Australian policy. Through a desktop search and consultation with experts, five relevant frameworks were identified as most relevant to the Australian context (see Table 1).

Table 1: Relevant education quality frameworks and evaluation tools

Quality Frameworks
Framework for Improving School Outcomes (FISO) The FISO forms the basis of the school Improvement Model of the Department of Education and Training in Victoria. It is a multi-level framework consisting of eight 'Essential Elements' encompassing 16 'Dimensions', each of which contain multiple descriptions of school quality. There are also four state-wide 'Priority Areas' that link to Essential Elements. The framework is intended as a self-assessment tool for the school leadership group to guide school improvement efforts.
National School Improvement Tool (NSIT) The NSIT was developed by the Australian Council for Educational Research in collaboration with the Queensland Department of Education, Training and Employment, and was endorsed for use nationally by the Australian federal government in 2012. Using the tool, schools can judge their performance on the nine inter-related domains as either 'Low', 'Medium', 'High' or 'Outstanding'.
School Excellence Framework (SEF) The SEF was developed by the New South Wales Department of Education, drawing on research evidence and the NSIT. The framework is intended to help schools plan and monitor improvement processes. It comprises three quality practice domains (Teaching, Learning, and Leading) covering 51 'Themes' on which schools can judge performance as 'Working towards Delivering', 'Delivering', 'Sustaining & Growing' or 'Excelling'. Self-assessments are conducted annually and externally validated every 5 years.
National Quality Framework (NQF) The benchmarking framework of Australian Children's Education & Care Quality Authority (ACECQA) is nationally implemented across the ECEC sector. Services receive a publicly available rating on each of the seven Quality Areas. Since 2015 the NQF framework has also been implemented in Western Australian schools up to Year 2.
Standards in Education, Children's Services and Skills (Ofsted) school inspection framework This framework guides school quality and effectiveness assessment in England. State-funded schools and some independent schools are assessed approximately every 2-3 years and their performance is rated across 5 areas as either 'Outstanding', 'Good', 'Requires Improvement' or 'Inadequate'.

The domains, dimensions, quality standards, and elements of these frameworks (FISO, NSIT, NQF and Ofsted respectively) were then examined with a view to identify common themes (see Appendix A).

Literature search

Search Strategy Overview

Systematic reviews and meta-analyses constitute the highest levels of evidence according to the National Health and Medical Research Council hierarchy of evidence (NHMRC, 2009). Accordingly, we conducted a broad search for systematic reviews and meta-analyses by searching relevant academic databases. The search was conducted in two phases. In the first phase, search terms were kept broad and papers published from January 2012[4] onwards were eligible for inclusion. In contrast, the second phase of the search used domain-specific search terms and a wider publication date range (from 1990 for meta-analyses and from 2008 for randomised controlled trials[5]).

Phase 1 Search Strategy

The following databases were used to identify relevant primary literature: ERIC, PsychINFO, Cochrane library, Medline, A+ Education, and ProQuest education database. As noted above, the search terms in the first phase of searching were kept broad. The Title/s, Abstract/s, MeSH terms, and Keywords lists were:

- *early childhood, primary or elementary*
- *school, education, classroom, class or learning*
- *systematic review, meta-analysis, or metaanalytic*

See Appendix B for a complete record of the searches conducted.

Systematic reviews or meta-analyses were included if they evaluated any aspect of schools related to quality, and if they satisfied the criteria outlined in Table 2.

Phase 2 Search Strategy

The second phase of the literature search was informed by the existing frameworks of school quality and findings from the first phase of searching. The frameworks provided the context from which a targeted search strategy was developed and themes arising from systematic reviews and meta-analyses were also incorporated in the targeted search strategy.

In Phase 2, separate searches were conducted for each domain, tailored according to the level of evidence already identified in Phase 1. A search for meta-analytic evidence was indicated for most domains. However, in some instances, the findings from Phase 1 indicated that a search at the level of randomised controlled trials would be more appropriate (see Appendix C)

The search strategy for meta-analyses had the form: (meta-analysis terms) AND (specific strategy terms). Keywords were searched in titles or abstracts. The search strategy for RCT level evidence had the form: (early childhood terms) AND (school-based intervention terms) AND (strategy-specific terms) AND (study design terms). Database specific subject heading searches were used where they added value. Database limiters such as population age were explored but not applied as the number of records returned was manageable without imposing such restrictions.

[4] Exceptions were made for meta-analyses considered both highly relevant and seminal in education research (e.g. #100)

[5] For domains where the volume of papers included at full-text screening was large, further date restrictions were applied (details are presented in the Results section).

In Phase 2, only peer-reviewed manuscripts published in academic journals, in the English language, were accepted. Databases consulted included ERIC, Education Resource Complete, PsycINFO, and Cochrane.

Paper Selection and Quality Assessment

Results obtained in both Phase 1 and Phase 2 searches were screened on abstract, title, keywords, and inspection of full text according to the criteria outlined in Table 2.

Table 2: Screening criteria for systematic reviews, meta-analyses and RCTS.

Criterion	Description
Population Focus	• Systematic reviews and meta-analyses had to include studies of children in primary school grades[6]. • Randomised controlled trials had to focus on children in the early years of primary school (i.e. Foundation[7] to Grade 3, typical age range 4 to 9 years). • In the second phase of searches, an additional population criterion was introduced for domains supported by an extensive body of literature. In the curriculum and digital technology domains meta-analyses were restricted to those presenting pooled effect sizes specific to the early years of primary school (i.e. grades K-3). Similarly, in the peer teaching domain, meta-analyses were included only if they presented pooled effect sizes specific to the early years or to primary school age students. • Reviews, meta-analyses and RCTs focussing on children with very specific health conditions or disabilities (e.g. autism) were excluded (studies of children with learning difficulties or disorders who attend a mainstream educational setting were included).
Intervention	• Systematic reviews and meta-analyses had to focus on school-based interventions, or present data for school-based interventions separately. • To be considered school-based, interventions had to be delivered on school grounds, during normal school staff hours, by 'regular' school staff (i.e., personnel schools can typically afford). • Studies of summer holiday programs were excluded.
Comparison group	• To be included studies had to include a comparison group; usual practice, alternative treatment, or waitlist control comparison group.
Outcomes	• In the first search phase, the study had to present evidence of an association between a school quality domain and at least one cognitive, social, emotional or behavioural student outcome. • In Phase 2 searches, relevant outcomes were domain-specific (see Results for details).
Study Design	• In the first search phase, the review or meta-analysis had to include quantitative studies. Reviews of purely qualitative studies and validation studies were excluded. • In the second search phase, where domain-specific strategies were targeted, meta-analyses had to include experimental or quasi-experimental group-level studies. Searches for individual studies were restricted to RCTs or cluster RCTs. However, owing to a paucity of

[6] If the focus was explicitly on children in grades 4 or above, the paper was excluded.
[7] In US-based studies, the first year of school is Kinder (children are typically 5 years of age). We therefore included studies of kindergarten children.

Criterion	Description
	research in the physical environment domains, within-participant counterbalanced repeated measures experiments in artificial classroom settings were included. For the same reason, quasi-experimental designs were included for the external collaborations domain.
Time Frame	• In the first search phase, documents published prior to 2012 were excluded. • In the domain-specific targeted searches, meta-analyses were generally excluded if published prior to 1990. • For two domains a tighter publication timeframe was applied (2010 for digital technology and curriculum research, owing to the importance of recency and volume of research, respectively). • In RCT-level searches studies published prior to 2008 were excluded.
Language	• Publications in non-English languages were excluded. Studies of interventions primarily delivered in languages other than English were also excluded.
Country	• Reviews and meta-analyses primarily focusing on studies conducted in low and middle-income countries were excluded. • RCTs conducted in low or middle-income countries were excluded.
Publication Type	• The research had to be published in an academic journal, or on a reputable evidence database website such as the Best Evidence Encyclopedia. • Reports were included in the first search phase but excluded in the domain-specific searches. • Dissertations, books, and book chapters were excluded.

Following title and abstract screening a sample of 15% was randomly selected and double-checked by a second reviewer (CM). Overall agreement was high (ranging from 85%-100% across domains, mean/median 94%), and all discrepancies were resolved. Full text versions of papers meeting eligibility criteria at the title and abstract stage were then sourced for screening of Method and Results sections by another qualified reviewer (RB). At this point 15% of the articles were again randomly selected and checked (CM). No discrepancies were identified. Papers meeting the inclusion criteria were subject to data extraction.

Quality and bias assessments were conducted for all publications included following full-text screening. The quality of systematic reviews and meta-analyses was assessed using the PRISMA checklist (Moher et al. 2009) (see Appendix D). RCTs and cluster RCTs were appraised for quality and bias using the National Institute for Health and Care Excellence (NICE) checklist (Appendix E). These checklists were completed by trained research assistants and a random sample of 15% was independently double-coded to check reliability.

Synthesising & Ranking the Evidence

The next step was to synthesise and appraise the evidence identified in the literature relating to each of the potential school quality domains. For each quality domain, we first examined the relevant evidence arising from systematic reviews or meta-analyses and extracted the relevant data.

Data extraction codes were constructed and organised according to the following broad categories: study characteristics (e.g. meta-analysis or systematic review, number of studies, included designs, publication date range), sample (e.g. age or grade range, countries, sample size), interventions (e.g. type of strategy, type of control groups, intervention durations, targeted or universal delivery), child outcomes (e.g. academic, social, emotional, behavioural), findings (e.g. effect sizes or conclusions, maximum length of follow-up), quality rating (PRISMA/NICE score), and proportion relevant (e.g. percent of studies focusing on primary or lower primary age students).

RCTs were summarised in terms of the study design, setting, intervention, sample and findings.

An overall appraisal of the evidence linking specific strategies within each of the quality domains to child cognitive, social, emotional and behavioural outcomes was then conducted according to the following factors:

1. Strength (quantity and type) of the evidence
    - Consistent with the NHMRC hierarchy of evidence, findings of systematic reviews and meta-analyses were considered as the strongest form of evidence, followed by findings from randomised controlled trials
    - The number of studies and sample sizes were considered to determine the quantity of evidence
2. Consistency – across studies, populations, & study designs
3. Generalisability & applicability – how well findings can be generalised to the Australian schooling context, given the participants & settings from which the existing evidence is drawn

The strength of the evidence supporting strategies identified in each quality domain was then ranked. Strategies with the strongest level of evidence were rated Well Supported. These were followed by Supported and then Promising strategies. The evidence base for all other strategies could be described as Preliminary, Unknown, Not Effective, or Concerning. More detailed descriptions of the evidence ranking scheme are presented in

Table **3**. To ensure evidence rankings were reliable, all were double-coded. Overall agreement was high (95%), with all discrepancies resolved.

Finally, an overall evidence ranking was developed for each *domain*. The same terminology was used for the evidence levels, but assessments were made *across* strategies for each domain (see Appendix F for the strategy and domain level evidence ranking systems). For example, Well Supported domains comprised at least two meta-analyses or systematic reviews identifying different strategies rated Well Supported. Supported domains comprised at least one meta-analysis or systematic review identifying a Supported or Well Supported strategy. To receive a Promising rating there had to be at least two high quality RCTs identifying different types of strategies with demonstrated effectiveness. Where there was only high quality quasi-experimental study or moderate quality RCT identifying an effective strategy, the evidence level for the domain was rated Preliminary.

Table 3: Study-level evidence rating system

Overall ranking of the evidence-study level	
Level of Evidence	**Definition**
Well Supported	Clear, consistent evidence of benefit. No evidence of harm or risk to participants. A well conducted[8] systematic review or meta-analysis found the intervention to be more effective than a control group on at least one child valid outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). Populations examined are similar to, and results are sensible to apply to, the Australian primary[9] school context.
Supported	Evidence suggestive of benefit but more evidence needed. No evidence of harm or risk to participants. A systematic review or meta-analysis of moderate quality[10] found the intervention to be more effective than a control group on at least one child valid outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the review are sensible to apply to primary school age students. Populations examined may be somewhat different to the Australian population; affecting generalisability to the Australian context.
Promising	No evidence of harm or risk to participants. At least one RCT with low to moderate risk of bias found the intervention to be more effective than a control group on at least one valid child outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the study are sensible to apply to primary school age children[11], though populations may be somewhat different to the Australian population.
Preliminary	No evidence of harm or risk to participants. At least one QES with low risk of bias found the intervention to be more effective than a control group on at least one valid child outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the study are sensible to apply to primary school age children[12], though populations may be somewhat different to the Australian population.
Not supported	A well conducted systematic review or meta-analysis or at least one RCT found the intervention to be ineffective across several primary outcomes

[8] To be considered well-conducted, meta-analyses and systematic reviews had to receive a PRISMA rating indicating low risk of bias (++) and at least 50% of included studies had to be RCTs, QESs, or matched comparison designs.
[9] For meta-analyses and systematic reviews to be considered relevant to the early years of school, at least 50% of included studies had to involve elementary school students or results reported separately for elementary students.
[10] Moderate quality means the meta-analysis or review received a PRISMA rating indicating moderate risk of bias (+) and included at least 50% RCT, QES, or matched-comparison designs.
[11] At least 50% of participants, or the average age of participants, must be within the primary school range (i.e. 4 years to 12 years).
[12] At least 50% of participants, or the average age of participants, must be within the primary school range (i.e. 4 years to 12 years).

Overall ranking of the evidence-study level	
Level of Evidence	Definition
	compared with a control group. The overall weight of the evidence does not support the benefit of the practice.
Concerning practice	A well conducted systematic review or meta-analysis reported that the direction of effects was undesirable across several outcomes. At least 1 RCT with low risk of bias showed the practice to have a negative effect.
Unknown	The intervention has not been adequately assessed. Available meta-analyses, reviews, or RCTs are limited either in terms of quality (low PRISMA/NICE rating) or relevance (to primary school age population).

Generalisability of the evidence

The likelihood that results may generalise to the early years of school was rated either Possible, Plausible or Very Plausible. In cases where it was unclear what proportion of studies specific to a well-supported or supported strategy involved studies of children in primary school, the likelihood of generalisation was rated Possible. In cases where the evidence for a specific strategy was based on a high proportion of studies with primary school students (i.e. at least 75%), the likelihood of generalisation was rated Plausible. Where the likelihood that strategies would generalise to the early years of school was rated Very Plausible, the evidence the supporting evidence for the specific strategy was based on either (a) a high proportion of studies focusing on students in the early years of primary school, or (b) results presented separately for studies of children in the early years of school.

Development of draft Indicators

A list of potential school quality domains was constructed based on: (a) themes in existing frameworks of school quality (FISO, NSIT, Ofsted and NQF) and (b) additional themes emerging from the literature. The strength and applicability of evidence supporting the various strategies identified within each of the school quality domains was then used to develop a shortlist of evidence-based school quality process indicators.

Expert Evaluation of draft Indicators

The list of indicators was vetted by Dr John Ainley, Principal Research Fellow, Australian Council for Educational Research. Dr Ainley was asked to independently comment on the developed list of school quality indicators.

RESULTS

The process of mapping school quality frameworks against each other and examining themes from the first phase of the literature search resulted in a total of 13 potential quality domains. These are listed in Table 4. Overall, the various domains can be categorised according to three overarching themes that loosely correspond to what might be described as the ecological layers of the school: (a) effective classroom pedagogical practices, (b) the school environment more broadly, and (c) providers and partnerships with parents and the wider community.

Table 4: Overview of domains

Domains	Primary Outcomes
Part I: Effective Classroom Pedagogical Practices (academic interventions)	
1. Application of pedagogical content knowledge	Student academic achievement & academic engagement (e.g. on-task behaviour)
2. Effective differentiated teaching	
3. Peer tutoring and collaborative learning	
4. Physical activity for academic achievement	
5. Technology-assisted teaching and learning	
6. Physical environment design to optimise learning	
7. Class size and Teacher-Student ratios	
Part II: School Environment and Student Wellbeing	
8. Student empowerment and leadership	Student social-emotional or behavioural outcomes (including school engagement) & staff-student relationships
9. Social-emotional and behavioural (SEB) interventions to promote a positive school climate	
10. Teacher-student relationships	
Part III: Providers and Partnerships (teacher & principal professional development, family engagement, community collaboration)	
11. Staff and leadership development	Student academic, social-emotional /behavioural and health outcomes
12. Partnerships with families	
13. Community-school partnerships	

The Phase 1 search produced 1070 results, of which 33 met inclusion criteria. Overall, Phase 2 searches identified a total of 3031 publications. Of these, 316 were included at title and abstract screening, and 50 at full text. In total, 83 studies were included. A flow diagram of the number of publications identified (and included at each stage of screening) in Phase 2 searches is presented in Appendix G; results by domain are presented in Appendix H. For a full list of included publications is see Appendix I.

Table 5 shows the overall evidence rating applied to each domain, together with the applicability of findings to the early years of school. It is important to note there was no evidence that strategies rated Supported or Well Supported in one publication were contraindicated in others. In cases where a similar strategy was Not Supported this was typically due to a low proportion of primary student studies.

Table 5: Summary of the evidence base by domain and applicability to the early years of school

Domains	Overall Evidence Rating*	Applicability to the Early Years of School
Part I: Effective Classroom Pedagogical Practices (academic interventions)		
1. Application of pedagogical content knowledge	Well Supported	High
2. Effective differentiated teaching strategies	Well Supported	High
3. Peer teaching and co-operative learning approaches	Supported	Moderate
4. Use of physical activity	Supported	Moderate
5. Technology-assisted teaching and learning	Well Supported	High
6. Physical environment design to optimise learning	Preliminary	High
7. Class size and Teacher-Student ratios	Supported	High
Part II: School Environment and Student Wellbeing		
8. Student empowerment and leadership	Unknown	Unknown
9. Social-emotional and behavioural interventions to promote a positive school climate	Well Supported	High
10. Teacher-student relationships	Preliminary	Moderate
Part III: Providers and Partnerships		
11. Staff and leadership development	Well Supported	High
12. Partnerships with families	Supported	High
13. Community-school partnerships	Preliminary	High

*Overall evidence ratings do not take into account applicability of evidence to EYS

PART I: EFFECTIVE CLASSROOM TEACHING PRACTICES

This section focuses on the literature relevant to effective classroom teaching practices. It includes approaches to teaching categorised into seven distinct domains. These include strategies that relate to the use of:

a) Application of pedagogical content knowledge,
b) Effective differentiated teaching,
c) Peer tutoring and collaborative learning,
d) Physical activity for academic achievement,
e) Technology assisted teaching and learning
f) Class size or student-teacher ratios.

Domain definitions and specific research questions guide each section.

Domain 1: Application of pedagogical content knowledge

The definition, of pedagogical content knowledge adopted in this report follows traditional formulations and refers to an understanding of instructional strategies to make specific subjects comprehensible. The focus is on both content-specific teaching strategies and the application of general instructional strategies to specific topics or skills being taught. For example, broad teaching strategies such as 'direct instruction' will only be included when tied to specific content or skills.

Research Questions:

1.1 Is there evidence to support the importance of applying pedagogical content knowledge?

1.2 When teaching specific subjects, which pedagogical strategies have demonstrated positive effects for academic achievement in the early years of school?

Strength of the domain

Domain 1 was rated Well Supported. Overall, the evidence level was high and consistent. Strategies were rated Well Supported or Supported in 11 of 13 publications. Moreover, the evidence shows seven strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

13 meta-analyses and systematic reviews were identified (see Table 6 for a summary and Appendix J for details). Of these, the evidence was rated Well Supported for two publications, and Supported for nine others. The types of strategies for applying pedagogical content knowledge and the extent to which the evidence is relevant to the early years of school is summarised next.

Table 6: Summary of meta-analyses and systematic reviews (Pedagogical content knowledge)

Study	Design	K-Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Abrami et al. 2015)	MA	684	K-12	Several instructional strategies [a]	Critical thinking	0.39* (#ES=867)	0.37*, #ES=49 (age 6-10 yrs)	High	Supported
(Carbonneau, Marley, and Selig 2013)	MA	55	K-12+	Manipulatives	Mathematics	0.37* (k=55)	0.33*, k=10 (age 3-6 yrs)	Moderate	Supported
(Elleman 2017)	MA	25	K-12	Inference instruction	Reading Comprehension	0.58* (k=13)	NR (82% primary)	High	Supported
(Goodwin and Ahn 2013)	MA	30	PK-12	Morphological Instruction	Literacy	0.32* (k=92)	0.68*, k=17 (grades PK-2)	Moderate	Supported
(Gordon, Fehd, and McCandliss 2015)	MA	13	K-4	Music instruction	Literacy & language	0.20* (k=13)	NR (77% EYS)	High	Supported
(Graham et al. 2018)	MA	47	PK-12	Balanced reading & writing	Reading Writing		0.46*, k=7 0.33*, k=6 (grades PK-1)	High	Well Supported
(Graham and Santangelo 2014)	MA	53	PK-12	Direct instruction	Spelling Phonological awareness Reading	0.54* (k=25) 0.51* (k=7) 0.44* (k=20)	0.63*, k=NR NR NR (grades K-2)	High	Supported
(Hammill and Swanson 2006)	MA[b]	38	K-6	Phonics instruction	Decoding Reading Comprehension Spelling	0.60*-0.67* #ES 30-40 0.25* (#ES=16) 0.27* (#ES=35) 0.35* (#ES=7)	Overall [c]: 0.55*, #ES=30 (grades K-1)	Moderate	Supported
(Piasta and Wagner 2010)	MA	68	PK-6	Phonics & phonological awareness instruction	Literacy skills [d]	See EYS	0.43 to 0.65 (83% PK-1)	High	Well Supported
(Santangelo and Graham 2016)	MA	76 (25 rel)	K-12	Handwriting instruction	Legibility Fluency	0.59* (k=20) 0.63* (k=15)	NR	High	Supported

Study	Design	K-Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
							(Pr:84%; 52% EYS of relevant studies)		
(Slavin et al. 2012)	SR	17	K-6	Science instruction	Science learning	0.02	NR	Moderate	Not Supported
(Stockard et al. 2018)	MA	328	K-4	Direct instruction	Reading Numeracy Language Spelling	0.51* (k=226) 0.55* (k=70) 0.54* (k=56) 0.66 *(k=52)	NR 0.17* NR 0.37* (grades PK-2)	Low	Unknown
(Suggate 2016)	MA	16	PK-6	Various [e]	Literacy [f] Post test Follow-up	0.37 (k=71) 0.22 (k=71)	0.34-0.40, k~30 0.12-0.26,k~30 (grades PK-2)	Moderate	Supported

[a] Instructional strategies such as use of dialogue, authentic and anchored instruction, and mentoring; [b] reanalysis of Ehri 2001; [c] See Appendix J for specific outcome effects; [d] letter-sound knowledge and fluency, letter name knowledge and fluency and letter writing; [e] phonemic awareness, phonics, fluency, reading comprehension, mixed interventions; [f] literacy skills, reading, comprehension, spelling composite; *p<0.05; rel= number of studies relevant

A variety of specific strategies derived from pedagogical content knowledge were identified and applied to the delivery of literacy, numeracy, handwriting, and science content. Well Supported instructional strategies included literacy programs that balance reading and writing instruction time (Graham et al. 2018), and phonics and phonological awareness instruction (Piasta and Wagner 2010). Supported instructional strategies included a number of different approaches to improve critical thinking (Abrami et al. 2015), direct teaching of spelling skills (Graham and Santangelo 2014), provision of explicit handwriting instruction (Santangelo and Graham 2016), inference instruction for reading (Elleman 2017), music education to benefit reading skills (Gordon, Fehd, and McCandliss 2015), various reading interventions (Suggate 2016), phonics instruction (Hammill and Swanson 2006), morphological instruction (Goodwin and Ahn 2013) and use of manipulatives in mathematics instruction (Carbonneau, Marley, and Selig 2013). The supporting evidence for each of these is briefly described below.

Table 7 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 7: Effective pedagogical content knowledge strategies and EYS applicability

Overall evidence rating	Observed magnitude of effect	Likelihood that effects generalise to early years of school*
Well Supported strategies		
• Balanced reading and writing instruction (Graham et al. 2018)	Small to moderate	Very Plausible
• Phonics and phonological awareness (Piasta 2010)	Moderate to large	Very Plausible
Supported strategies		
• Use of dialogue, authentic problem-solving, and mentoring in teaching to improve critical thinking (Abrami et al. 2015)	Small to moderate	Plausible
• Direct teaching of spelling skills (Graham and Santangelo 2014)	Small to moderate	Very Plausible
• Use of manipulatives in mathematics instruction (Carbonneau, Marley, and Selig 2013)	Small to moderate	Very Plausible
• Provision of explicit handwriting instruction (Santangelo and Graham 2016)	Moderate	Plausible
• Inference instruction for reading comprehension (Elleman 2017)	Moderate	Plausible
• Phonemic awareness, phonics, fluency, and reading comprehension (Hammill and Swanson 2006; Suggate 2016)	Small to moderate	Very Plausible
• Morphological instruction (Goodwin and Ahn 2013)	Moderate to large	Very Plausible
• Music instruction (Gordon, Fehd, and McCandliss 2015)	Small	Very Plausible

* Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%). Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small to Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate to Large=0.60 to 0.79; Large>0.80

Strategies rated Very Plausible

Well Supported and Supported subject-specific instructional strategies with positive effects considered very plausible to apply to the early years of school included:

- Phonemic awareness, phonics, fluency, and reading comprehension
- Balanced reading and writing instruction
- Morphological instruction (Goodwin and Ahn 2013)(Goodwin and Ahn 2013)(Goodwin and Ahn 2013)(Goodwin and Ahn 2013)(Goodwin and Ahn 2013)(Goodwin and Ahn 2013)
- Direct teaching of spelling skills
- Music instruction to teach reading skills
- Use of manipulatives in mathematics instruction
- Specific instructional strategies to teach critical thinking

The evidence for each of these is summarised below.

Reading: Phonemic awareness, phonics, fluency, and reading comprehension

Three publications identified instructional reading strategies with a focus on phonemic awareness and phonics; all demonstrated positive effects on reading or early literacy skills and were highly relevant to children in the early years of school.

The first meta-analysis evaluated the effects of targeting phonemic awareness, phonics, fluency, and comprehension (Suggate 2016). It included 16 studies of children from grades Pre-K to 6 and reported small to moderate positive effects at post-test (ES=0.37) and follow-up (ES=0.22, average follow-up time was about 11 months). Importantly, effects were reported specifically for children in the early years of school. For children in grades Pre-K to 2, positive effects of small to moderate magnitude were observed for interventions targeting phonemic awareness, phonics knowledge, fluency, and comprehension at both post-test (ES=0.34 to 0.40) and follow-up (ES=0.12 to 0.26). Overall, the evidence for phonemic awareness, phonics, fluency, and reading comprehension was rated Supported rather than Well Supported because there was moderate risk of bias and the countries from which studies were drawn these were not adequately reported.

A second meta-analysis (Piasta and Wagner 2010) included 63 studies of interventions to improve alphabet knowledge, many of which were multi-component interventions, with most including alphabet instruction and phonological training. It reported moderate and moderate to large effects on several alphabet outcomes, including letter sound knowledge (ES=0.65), letter sound fluency (ES=0.58), letter name knowledge (ES=0.43), and letter writing (ES=0.59). Importantly, the majority of effect sizes were calculated from studies of children in grades Pre-K to 1 (>75%). The evidence was rated Well Supported because the study reported positive effects of the included interventions, had low risk of bias, and countries from which studies were drawn were considered reasonably similar to the Australian school context.

A third meta-analysis evaluated the effect of phonics instruction for students in grades K to 6 (Hammill and Swanson 2006). This analysis (re)examined results from 38 experimental and quasi-experimental studies in the National Reading Panel report (Ehri et al. 2001). Large positive effects were reported for some literacy skills (e.g. decoding), with smaller positive effects on others (such as oral reading, comprehension, and spelling). Among children in grades K to 1, there was a demonstrable positive effect of moderate magnitude on literacy skills overall (ES=0.55). The evidence for phonics instruction

was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn.

Summary

Overall, the likelihood that the positive effects of instructional reading strategies with a focus on phonemic awareness and phonics would generalise to the early years of school was rated very plausible; this is because each of the three publications presented a pooled effect size specific to children in the Pre-K to grade 2 range.

Balanced reading and writing instruction

Literacy interventions that balance the amount of reading and writing instruction time (i.e. no more than 60% of instructional time allocated to either reading or writing) were evaluated in a well-conducted meta-analysis (Graham et al. 2018). This meta-analysis of 47 RCT and quasi-experimental studies included children from grades Pre-K to 12 and reported several small to moderate positive effects for both reading and writing outcomes. Indeed, positive effects of balanced instructional programs were observed on measures of decoding, fluency, and reading comprehension, as well as writing quality, writing mechanics and writing output. Of note, seven studies tested whether a balanced reading/writing program implemented with students in Pre-K to grade 1 enhanced total reading performance. The overall effect for these was statistically significant and approaching moderate magnitude (ES=0.46). A similar result was found for writing performance, which included six studies (ES=0.33) for students from grades Pre-K to 1. The evidence for balanced reading and writing instruction was rated Well-Supported because the meta-analysis was well-conducted, identified positive effects on academic outcomes, included a large proportion of primary age students, and was based on studies conducted in countries considered similar to Australia. The likelihood that the positive effects of balanced instruction would generalise to the early years of school was rated Very Plausible because the meta-analysis presented a pooled effect size specific to children in the early years of school.

Music instruction to teach reading skills

Teaching phonological and literacy skills within a musical context can involve a variety of strategies. Examples include the use of chants, songs, rhymes, and instruments to teach rhyming skills, letter sounds, or vocabulary. A meta-analysis including 13 experimental and quasi-experimental studies of children in grades K to 4 reported a small positive effect for phonological awareness (ES=0.20). Although effects were not reported separately by grade level, 77% of studies included children in grades K to 3 and most were conducted with typically developing children.

The evidence for music instruction was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn, and information about the languages spoken for each indicate that several studies were conducted in countries with education systems that might not be comparable to Australia's education system. The likelihood that the positive effects of music instruction would generalise to the early years of school was rated Very Plausible because a large proportion of studies included in the meta-analysis focused on children in the early years of school.

Although the evidence for music training is rated Supported, important limitations should be noted. First, it is unclear what proportion of music instruction interventions were implemented by classroom teachers. While the paper does indicate that there were cases where music instruction was implemented at the classroom or school level, it gives little information about how many interventions

were school-based or who implemented them. Second, further research is needed to determine how much music training is necessary to generate a practical effect, and which components of music instruction (e.g. rhyming, clapping, singing, movement, notation) are most effective.

Direct teaching of spelling

There was one meta-analysis of various instructional strategies to teaching spelling (Graham and Santangelo 2014). This evaluation included 53 RCT and quasi-experimental studies of students from K to year 12. Positive effects of moderate magnitude were reported for interventions where spelling instruction was directly compared with either no instruction or informal/incidental instruction. Large improvements favouring intervention groups were also observed on measures of spelling while writing, and moderate effects emerged on phonological awareness and reading measures. A statistically significant effect of moderate magnitude also indicated a positive effect on spelling (ES=0.63), in studies of children in grades K to 2. As such, the likelihood that positive effects would generalise to the early years of school was rated Very Plausible. The evidence for direct teaching of spelling was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn.

Morphological instruction

One meta-analysis evaluated morphological instruction interventions for literacy outcomes (Goodwin and Ahn 2013). Morphological interventions are those which teach students to identify and analyse units of meaning (i.e., roots and affixes) within words to support literacy development. The meta-analysis of 30 studies including experimental and quasi-experimental studies reported statistically significant positive effects of small to moderate magnitude on several literacy outcomes including decoding (ES=0.59), phonological awareness (ES=0.48), morphological knowledge (ES=0.44), vocabulary (ES=0.34), and spelling (ES=0.30). Results specific to children in grades K-2 indicated a moderate to large effect on literacy measures overall (ES=0.68, k=17). Therefore, the effects of morphological instruction were considered very plausible to apply to the early years of school. The evidence was rated Supported rather than Well Supported because the meta-analysis had a moderate risk of bias and did not report the countries from which studies were drawn.

Mathematics: Use of manipulatives

Manipulative-based instruction is defined as teaching strategies that provide students with opportunities to physically interact with objects to aid learning. Examples of manipulatives include play money, counting blocks, and Cuisenaire rods (not use of rulers, scales or calculators). One meta-analysis (Carbonneau, Marley, and Selig 2013) investigated the use of manipulatives in 55 studies of mathematics instruction, including 13 experiments and 30 quasi-experimental designs. This meta-analysis reported positive effects of moderate magnitude across a range of mathematical topics including place values, arithmetic, geometry, fractions, and algebra. With regard to child grade, the use of manipulatives (compared with teaching mathematical concepts using abstract representation) resulted in positive effects for children aged 7 to 11 years (ES=0.45), and children ages 3-6 years (ES=0.33). Positive effects were therefore considered Very Plausible to apply to the early years of school. The evidence for use of manipulatives in maths instruction was rated Supported rather than Well Supported because the meta-analysis had moderate risk of bias, and did not report the countries from which the studies were drawn.

Instructional strategies to critical thinking

The evidence for instructional critical thinking interventions is based on one meta-analysis of 341 effect sizes drawn from experimental and quasi-experimental investigations using standardised measures of critical thinking outcomes (Abrami et al. 2015). Several effective strategies for teaching critical thinking

including creating opportunity for dialogue, exposing students to authentic or situated problems and examples, and mentoring were identified. Specific dialogue strategies where small to moderate effects significantly favoured experimental groups over comparison groups included: teachers posing questions, whole-class discussions led by the teacher, and small-group discussions led by the teacher (ESs=0.38 to 0.42). Applied problem-solving and role-playing were identified as effective subcategories of authentic and anchored instruction. Mentoring was effective, though there was no evidence that specific sub-categories of mentoring were more effective than others. The overall effect of different strategies for teaching critical thinking to students (ranging from grades K to 12+) was small to moderate (ES=0.30). The effect for children 6 to 10 years of age was similar (ES=0.37). Unfortunately, the effectiveness of *specific* strategies was not reported separately by age group, so it is unclear which specific strategies are most relevant to children during the early years of school.

The overall evidence for instructional critical thinking interventions was rated Supported rather than Well Supported because there was no information reported about the countries from which the studies were drawn. As an effect size specific to children in the range of 6 to 10 years was reported, the evidence was considered Very Plausible to apply to the early years of school.

Strategies rated Plausible

Inference instruction for reading comprehension

One meta-analysis evaluated inference instruction interventions for reading comprehension (Elleman 2017). Inference instruction interventions included a variety of strategies, such as teaching students to activate background knowledge, monitor comprehension, generate questions before and after reading, examine text for clues, and elaborate on textual information The meta-analysis of 25 RCT and quasi-experimental studies reported positive and statistically significant effects of moderate magnitude on measures of literal, inferential, and general comprehension (ESs= 0.28, 0.68, and 0.58, respectively) across students in grades K-12. As many studies contained multiple inference strategies, the analysis was unable to determine which were most effective. Nevertheless, the overall results show that combined inference instruction strategies are generally effective. As a large proportion of studies included primary-school age students, the extent to which results might generalise to children in the early years of school was considered plausible. The evidence for inference instruction strategies was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn.

Handwriting instruction

There is evidence from one meta-analysis (Santangelo and Graham 2016) of 76 RCTs and quasi-experimental studies that providing handwriting instruction (compared with no instruction) has a significant and positive effect of moderate magnitude on both handwriting fluency (ES=0.63) and legibility (ES=0.59). Positive effects of moderate to large magnitude were also observed from studies of handwriting instruction on writing quality (ES=0.84), length (ES=1.33), and fluency (ES=0.48). Although results were not reported separately for children in the early years of school, 73% of studies focused on children in the early years of primary school. The extent to which results are likely to generalise to the early years of school was therefore considered plausible.

In addition to investigating whether explicit handwriting instruction is effective, the meta-analysis also evaluated specific instructional methods characterising handwriting interventions. These included use

of self-evaluation as part of hand-writing instruction (vs a variety of control conditions[13]), teaching individual letters with motion models (vs still models), and copying letters from models (vs no instruction or no repetition). Although small positive effects on legibility were also observed for these methods the results were based on only a small number of studies (k=4-5) and failed to reach statistical significance. The evidence for handwriting instruction was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn.

Domain 1 Conclusions

The domain 'Application of pedagogical content knowledge' was rated Well Supported.

When delivering specific curriculum content, which instructional strategies have been shown to be most effective in the early years of school?

The specific strategies shown to be most effective in the early years of school are those targeting: phonemic awareness, phonics, fluency, and reading comprehension; balanced reading and writing instructional time; explicit morphological instruction; directly teaching spelling skills; use of musical context to teach reading and literacy skills; provision of explicit handwriting instruction and use of manipulatives in mathematics.

[13] Two control conditions involved the *Palmer Handwriting Program*, one included instruction with peer-evaluation, and the fourth used printed models for self-evaluation (as compared to transparent overlays in the intervention—which was considered to be a stronger form of self-evaluation).

Domain 2: Effective differentiated teaching

We define 'differentiated teaching' as modifications to instructional delivery that enable teachers to tailor instruction to the needs of students across a range of abilities and learning needs. The terms differentiated, tailored, and individualised instruction are considered synonymous in this report. However, the term 'individual teaching' is different from 'individualised' teaching. Individual tutoring and targeted small group programs delivered outside regular classroom hours or by providers other than regular classroom teachers are beyond the scope of this domain.[14]

Research Questions:

2.1 Is there evidence that differentiated teaching can improve student academic outcomes?

2.2 Which differentiated teaching strategies have demonstrated positive effects on academic achievement for children in the early years of school?

Strength of the domain

Domain 2 was rated Well Supported. Differentiated teaching strategies were rated Well Supported or Supported in four of six publications. Moreover, the evidence suggests two strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were six meta-analyses and systematic reviews identified (see Table 8) for a summary and Appendix J for details). Two strategies were rated Well Supported, two Supported and two Unknown. Findings are summarised in Table 9.

[14] Readers interested in small-group and 1:1 interventions are referred to Dennis (2016) for mathematics outcomes and Wanzek (2016) for reading outcomes.

Table 8: Summary of meta-analyses and systematic reviews (Differentiated teaching strategies)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Deunk et al. 2018)	MA	21	~K-6[a]	Various [b]	Language and mathematics composite	0.146* (k=21)	No pooled results[c]	High	Well Supported
(Graham, Hebert, and Harris 2015)	MA	35	K-8	Individualised feedback (various sources)[d]	Writing quality	0.61* (k=35)	NR (72% primary)	High	Supported
(Lou, Abrami, and Spence 2000)	MA	51	K-12+	Within-class grouping vs whole class instruction	Cognitive (NOS)	0.16* (103 ES)	NR (grade moderator, stronger earlier, effects NR)	Moderate	Unknown
(Ok et al. 2017)	SR	13	PK-12	Universal design for learning	Academic [e]	'small to large'	NR	Moderate	Unknown
(Piasta and Wagner 2010)	MA	63 (3-11 rel.)	PK-1	Group size as moderator for literacy instruction	Literacy skills: Letter name Letter sound Letter writing Letter name fluency	Small group vs whole class: 0.52*, k=21 vs. 0.24*, k=4 0.73, k=2 vs. 0.48, k=11 0.56, k=2 vs 0.60, k=3 0.07, k=7 vs 0.06, k=5		High	Well Supported
(Santangelo and Graham 2016)	MA	76 (8 rel.)	K-12	Individualised handwriting instruction	Legibility	0.69* (k=8)	NR (89% primary – this domain)	High	Supported

[a]Ages 6-12 years; [b] Homogenous ability grouping, computerised differentiation, and differentiation in broader reform context-well supported overall but NOT necessarily individually; [c] There were some positive effects for computerised differentiation; [d] Study examines effects for source of feedback: adults, peers, self, computer; [e] reading, science, social science studies; NOS= Not Otherwise Specified; *p<0.05

Several specific differentiated teaching strategies were identified. Well Supported strategies included homogenous ability grouping and computerised differentiation (Deunk et al. 2018) and instruction in small groups compared with whole-class instruction (Piasta and Wagner 2010). Supported strategies included providing individualised feedback from various sources (Graham, Hebert, and Harris 2015; Dennis et al. 2016; Santangelo and Graham 2016).

Table 9 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 9: Effective differentiated teaching strategies and EYS applicability

Overall evidence rating	Observed magnitude of effect	Likelihood that effects generalise to early years of school
Well Supported strategies		
• Small group instruction in literacy (Piasta and Wagner 2010)	Moderate to large	Very Plausible
• Within-class homogenous ability grouping (Deunk et al. 2018)	Mixed results from individual studies	Possible
• Computerised differentiation (Deunk et al. 2018)	Small	Very Plausible
Supported strategies		
• Individualised writing quality feedback from appropriately trained adults, peers, or self-assessment (Graham, Hebert, and Harris 2015)	Moderate to large	Plausible
• Provision of individualised handwriting instruction (Santangelo and Graham 2016)	Moderate to large	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are reported. Possible: proportion of studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Strategies rated Very Plausible

Two Well Supported or Supported differential teaching strategies with positive effects considered Very Plausible to apply to the early years of school were identified. These are described next:

Small group instruction in literacy skills

A meta-analysis of 63 studies investigating alphabet training effects on early literacy skills (such as letter name and sound knowledge) included 21 studies utilising small group instruction (Piasta and Wagner 2010). A positive moderate effect (ES=0.52) of small group instruction on improved letter name knowledge was found. In contrast, studies utilising whole-class instruction reported a smaller effect (ES=0.24) on improving letter name knowledge. These results suggest that presenting literary material in small groups is effective for children in grades Pre-K to 1.

The evidence for use of small group early literacy instruction was rated Well Supported and findings were considered Very Plausible to apply to the early years of school because all participants in the included studies were in grades Pre-K to 1.

Computerised differentiation

A meta-analysis and systematic review including 21 studies comparing various differentiation strategies with usual practice reported a positive effect for computerised systems of differentiation (ES=0.29) tested with primary school students (Deunk et al. 2018). Although results specific to the early years were not pooled, the review summarised three trials including only children in the early years of school. Each trial evaluated a reading intervention in which teachers were trained to use a computer program that made instructional suggestions based on student performance. Instructional suggestions included student grouping and selection of appropriate content. The first trial utilised a matched control group design and reported small but significant positive effects on reading achievement (ES=0.18) for grade 1 students. The second trial evaluated the same program with a different group of grade 1 students and also reported positive effects on word-reading scores (ES=0.25). Finally, a third trial tested the program with grade 3 students. Those receiving differentiated instruction from teachers trained to use the software demonstrated significantly higher reading comprehension scores (ES=0.19) than children in an alternative treatment condition.

The extent to which findings for computerised differentiated teaching apply to children in the early years of school was rated Very Plausible because results specific to children in grades 1 and 3 were presented (though not pooled). However, it should be noted that the three studies included in the review all tested the same program and may not generalise to other programs.

Strategies rated Plausible

Two Well Supported and Supported instructional strategies with positive effects considered Plausible to apply to the early years of school were identified. These included:

Individualised writing quality feedback

One meta-analysis of writing quality outcomes[15] including 35 studies reported effects of moderate-to-large magnitude (ES=0.61) for provision of individualised writing feedback (Graham, Hebert, and Harris 2015). Feedback from several different sources had positive effects. This included feedback from adults (ES=0.87), peers (ES=0.58), and self-assessment (ES=0.62). In most of these studies the person providing feedback received specific instructions or training in the use of a scoring rubric (i.e. appropriate training in provision of feedback was relatively common across these studies). Overall, the evidence base for providing individualised writing feedback was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn. Findings were rated Plausible to apply to the early years of school because the sample comprised a large proportion of primary students.

Individualised handwriting feedback

Evidence that individualised handwriting feedback is beneficial was identified in a meta-analysis of 76 studies (Santangelo and Graham 2016). Eight individualised handwriting studies yielded a positive effect of moderate magnitude on writing legibility (ES=0.69). Of the eight studies, four were conducted with children in grades K to 3. In half of the studies, individualised instruction was compared with no instruction and in the other half it was compared with non-individualised handwriting instruction. Although instructional strategies varied across the eight individualised interventions there was no detail provided about *how* the instruction was individualised. The lack of detail pertaining to the countries in which the interventions had been trialled resulted in an overall rating of Supported rather than Well Supported. Findings were rated Plausible to apply to the early years of school because a large

[15] Writing quality was defined in terms of reader judgements of overall merit (taking into consideration factors such as ideation, organisation, vocabulary, sentence structure etc.).

proportion of studies in the meta-analysis overall (and strategy-specific studies) included primary age students.

Domain 2 Conclusions

The domain differentiated teaching was rated Well Supported.

Which differentiated teaching strategies have demonstrated positive effects in the early years of school?

Specific differentiated teaching strategies shown to be effective in the early years of school are:

- Provision of literacy instruction in small groups.
- Computerised instructional differentiation.

Domain 3: Peer tutoring and collaborative learning strategies

Peer-tutoring strategies involve teachers training students to explain conceptual knowledge or provide structured feedback to other students (i.e. not merely modelling). Also often referred to as co-operative learning strategies, or peer assisted learning, peer tutoring strategies are considered structured, prescriptive, and directive. They may involve variations such as (a) reciprocal roles, (b) cross-age/ability or matched age/ability pairs or groupings, or (c) use of tangible or intangible rewards for appropriate behaviour or correct responses during peer tutoring lessons.

Collaborative learning strategies may also involve organising students in pairs or small mixed- or similar-ability groups to complete goal-directed activities involving verbal and social interaction. However, collaborative grouping is considered less structured, prescriptive or directive (Puzio 2013). Some conceptualisations of co-operative and collaborative learning make a distinction between students learning *from* (co-operative learning), as opposed to *with* (collaborative learning) one another. In this domain, we consider both co-operative and collaborative learning strategies.

Research Questions:

3.1 Is there evidence that peer-tutoring and collaborative learning strategies can improve student academic achievement?

3.2. Which specific peer-tutoring and collaborative grouping strategies have demonstrated positive effects on academic achievement for children in the early years of school?

Strength of the domain

Domain 3 was rated Supported. The evidence for peer tutoring and collaborative learning strategies was rated Supported in five meta-analyses and Well Supported in another. Moreover, the evidence shows the strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

Seven meta-analyses/systematic reviews were identified (see Table 10 for a summary and Appendix J for details). Strategies were rated Well Supported or Supported in six publications. Table 11 summarises the extent to which the evidence is applicable to the early years of school.

Table 10: Summary of meta-analyses and systematic reviews (Peer tutoring)

Study	Design	K Studies (# relevant)	Grades	Strategy [a]	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Dietrichson et al. 2017)	MA	101 (10 rel.)	K-9	Peer tutoring	Reading and Mathematics	0.22* (k=10)	NR	High	Unknown
(Elbaum et al. 1999)	MA	16	1-6	Peer tutoring (for students at risk of or with disabilities)	Reading outcomes	*glass Δ= 0.40* (#ES = 24)	NR	Moderate	Supported
(Kunsch, Jitendra, and Sood 2007)	MA	17	K-12	Peer tutoring	Mathematics	0.47 (k=17), p=NR	NR (but 0.57 Primary, k=14, and 3 individual studies presented)	Moderate	Supported
(Leung 2015)	MA	72	K-12+	Peer tutoring & collaborative learning	Academic (various)[b]	0.38* (overall, k=72) 0.47* (primary, k=46)	0.21 (kinder, k=2)	Moderate	Supported
(Puzio and Kolby 2013)	MA	18	2-12	Within-class peer tutoring & collaborative grouping	Reading Vocabulary Comprehension	0.16* (k=16) 0.20* (k=18) 0.22* (k=14)	NR (89% primary)	High	Well Supported
(Rohrbeck et al. 2003)	MA	90	1-6	Peer tutoring & collaborative learning	Academic (various)[c]	0.33* (k=81)	0.37* (grades 1-3, tutees, k=26)	Moderate	Supported
(Zeneli, Thurston, and Roseth 2016)	MA	41	K-12	Peer tutoring	Academic (various)[d]	0.25* (#ES=32)	NR (0.51* Primary, #ES 30)	Moderate	Supported

[a] We use the terms peer tutoring and collaborative learning as defined above, rather than the variety of terms that are used somewhat inconsistently throughout the literature; [b] Academic achievement included a range of subject areas (e.g. maths, reading, language, science and technology, physical education, psychology, education); [c] Academic achievement included a range of subjects: reading, mathematics, social studies, science, writing, language, literacy; [d] Achievement defined as maths, reading science or "other" outcomes (no examples for 'other' provided)

A variety of peer tutoring and collaborative learning strategies were identified. The evidence for both strategies was rated Well Supported in literacy applications (Puzio and Kolby 2013) and Supported across a variety of subject areas (Rohrbeck et al. 2003). Supported strategies also included peer tutoring for students at risk of, or experiencing, learning difficulties in reading or mathematics (Elbaum et al. 1999; Kunsch, Jitendra, and Sood 2007), and peer tutoring for students across a range of ability levels and subject areas (Leung 2015; Zeneli, Thurston, and Roseth 2016).

Table 11 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 11: Effective peer teaching strategies and EYS applicability

Overall evidence rating	Observed magnitude of effect	Likelihood that effects generalise to early years of school*
Well Supported strategies		
• Peer tutoring and collaborative learning approaches to teaching literacy skills (Puzio and Kolby 2013)	Small	Plausible
Supported strategies		
• Targeted peer tutoring in literacy and numeracy (Elbaum et al. 1999; Kunsch, Jitendra, and Sood 2007)	Moderate	Plausible
• Peer tutoring and collaborative learning across a range of student abilities and subject areas (Leung 2015; Zeneli, Thurston, and Roseth 2016)	Small to Moderate	Plausible
• Peer tutoring and collaborative learning across a range of subject areas (Rohrbeck et al. 2003)	Small to Moderate	Very Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%). Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear.

Strategies rated Very Plausible

Peer tutoring and collaborative learning

Taken together, peer tutoring and collaborative learning strategies were rated Very Plausible to apply to the early years of school across a range of subjects.

A variety of peer tutoring and collaborative learning strategies were investigated in a meta-analysis of 90 experimental and quasi-experimental studies (Rohrbeck et al. 2003). Studies spanned a range of subject areas (e.g. reading, mathematics, social studies, science, writing, language, literacy) for children in grades 1 to 6. Interventions included both dyadic and small group strategies, mostly same-age, cross-ability grouping strategies to match students with peer tutors, and most provided students with structured guidance around their roles and interactions (though a sizeable proportion, 42%, were considered unstructured). Most interventions involved reciprocal roles, and almost half of these included reward contingencies to promote reciprocity.

In addition to reporting an overall effect of peer tutoring and collaborative learning across several subject areas combined (ES=0.33), an analysis of results from the studies of children in grades 1 to 3

was reported. This analysis also revealed a statistically significant positive treatment effect of small to moderate magnitude (ES=0.37).

The overall analysis found larger effect sizes when interventions used interdependent reward contingencies, ipsative evaluation procedures, or provided students with more autonomy.

The evidence in this meta-analysis was rated Supported rather than Well Supported because there was moderate risk of bias and the countries from which studies were drawn was not reported. The findings were rated Very Plausible to apply to the early years of school because an effect size specific to children in grades 1 to 3 was reported.

Strategies rated Plausible

Well Supported and Supported peer tutoring and co-operative learning strategies with positive effects considered Plausible to apply to the early years of school were:

- Targeted peer tutoring in literacy and numeracy
- Peer tutoring across a range of abilities in a range of subjects
- Peer tutoring and collaborative grouping strategies to develop literacy skills

Targeted peer tutoring (for students with emerging or established learning difficulties)

Two meta-analyses investigated the use of peer tutoring targeting students at risk of (Elbaum et al. 1999) or experiencing (Kunsch, Jitendra, and Sood 2007) learning difficulties. The first meta-analysis investigated the use of different grouping formats in 20 reading intervention studies among grade 1-6 students with various disabilities (mostly learning disabilities, and/or dyslexia, behavioural difficulties, neurological impairment, or emotional disturbance). An effect of small to moderate magnitude was reported for interventions where the grouping format was in pairs (ES=0.40). A sub-group analysis of different types of pairs revealed small to moderate effects for both interventions utilising cross-age (ES=0.50) and same-age (ES=0.24) formats.

A second meta-analysis investigated peer tutoring interventions for children at risk of, or experiencing, difficulties in mathematics (Kunsch, Jitendra, and Sood 2007). In this meta-analysis the term peer mediated instruction was used, and it referred to the practice of pairing students to work collaboratively in structured, individualised activities. There were 17 experimental or quasi-experimental studies of students in grades K to 12. Peer tutoring had a positive effect of small to moderate magnitude (ES=0.47). Effect sizes were generally larger for studies of children considered at risk (ES=0.66) than those with established difficulties (ES=0.21). Importantly, the effect size obtained from the 14 studies of children in primary school grades was moderate (ES=0.50). Additionally, unpooled effects from three separate individual studies of children in grades K to 3 were also presented. One reported a positive effect for grade 2 students on teacher-graded mathematics (ES=0.12) and standard arithmetic measures (ES=1.45). Another reported positive effects of small to moderate magnitude on standardised achievement tests for students in kindergarten (ES=0.21 to 0.46).

The evidence supporting peer tutoring for children who are at risk of, or experiencing, learning difficulties was rated Supported because neither of the relevant meta-analyses reported the countries from which studies were drawn, and both were assessed as having moderate risk of bias. The findings were rated Plausible to apply to the early years of school because both meta-analyses reported effects specifically for primary school students, and one presented (unpooled) effects that were positive for two of three separate studies of students in the early years.

Peer tutoring and collaborative learning across a range of abilities and subject areas

Two meta-analyses investigated the use of peer tutoring with students characterised by a range of learning abilities (Leung 2015; Zeneli, Thurston, and Roseth 2016). The first investigated the use of peer tutoring and collaborative learning across 72 experimental and quasi-experimental studies covering a range of academic outcomes (e.g. maths, reading, language, science and technology, physical education, psychology, arts, education), for students from grades K to tertiary study. An overall positive effect of small to moderate magnitude was reported (ES=0.37). An effect of similar magnitude was also reported for children in primary school grades (ES=0.47), based on 46 individual studies. A smaller effect (ES=0.21) was reported for children in kindergarten but was based on only two studies.

Although not specific to children in primary school, this meta-analysis also found several interesting moderators of peer tutoring and collaborative learning effects. Specifically, results indicated that (a) structured tutoring interventions had larger effects than unstructured interventions (ES=0.53 vs 0.33), (b) interventions using tangible rewards had larger effects than those using symbolic rewards such as points (ES=0.70 vs 0.35), and (c) interventions involving same-gender dyads had larger effects than mixed gender dyads (ES=0.80 vs. 0.41).

The second meta-analysis focused on peer tutoring (Zeneli, Thurston, and Roseth 2016). It included students representing a range of learning abilities across 41 studies (32 of which were experimental or quasi-experimental). Studies explicitly focusing on learning-disabled students were excluded, and interventions were required to last at least six weeks or longer. It included children in primary and secondary school and defined peer tutoring as a specific form of co-operative learning where at least one student teaches another, either working in a pair or triad, with a clear pattern of interaction. Overall, the meta-analysis reported a small effect of peer tutoring (ES=0.25) measured across a range of academic subjects (e.g. reading, mathematics, social studies, science, writing, language). For primary school children specifically, a positive intervention effect of moderate magnitude was reported from experimental and quasi-experimental studies (ES=0.51) also incorporating a range of academic outcomes.

The evidence for peer tutoring across a range of student abilities was rated Supported because both meta-analyses had moderate risk of bias. The findings were rated Plausible to apply to the early years of school because effects specific to primary school children were reported, but effects specific to children in the early years of school were either not reported (or were based on too few studies to confidently generalise results).

Peer tutoring and collaborative learning for literacy skills

One meta-analysis (Puzio and Kolby 2013) included both studies of peer tutoring[16] and collaborative grouping to develop literacy skills (reading, vocabulary and comprehension). It included 18 experimental and quasi-experimental design studies with 29 cohorts spanning grades 2 through 12. Overall, small but statistically significant positive effects were reported on standardised measures (ES=0.16 to 0.22). Although results specific to the early years were not reported, the applicability was rated Plausible because a very high proportion of included studies (89%) were conducted with primary school students. The evidence rated Well Supported because there were positive effects, the analysis was well-conducted, and included studies were from countries with education systems considered similar to Australia (predominantly US-based).

[16] Note that 1:1 peer tutoring was excluded.

Domain 3 Conclusions

The domain peer tutoring and collaborative learning practices was rated Supported.

Which peer tutoring or collaborative learning approaches have demonstrated positive effects for students in the early years of school?

Overall, peer-tutoring and collaborative learning strategies have demonstrated positive effects in the early years of school. However, it is unclear which *specific strategies* are most effective for children in the early years.

Domain 4: Physical activity for academic achievement

This domain relates to teaching practices that utilise physical activity to improve academic outcomes. Physical activity is broadly defined as any bodily movement produced by the skeletal muscles that require energy expenditure (WHO, 2019). In the context of classroom approaches to teaching, the definition adopted here includes only movement that is under volitional control and involves more than maintenance of posture.

Research Questions:

4.1. Is there evidence that engaging students in physical activity can improve academic outcomes or school engagement?

4.2. Which types of physical activity have demonstrated positive effects on academic achievement and/or school engagement in the early years of school?

Strength of the domain

Domain 4 was rated Supported. The evidence that physical activity improves school engagement was rated Well Supported for one meta-analysis. Similarly, evidence that physical activity improves academic outcomes was rated Well Supported for one systematic review. However, the applicability of findings to the early years of school is limited to a few studies only. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were 2 meta-analyses/systematic reviews identified (see Table 12; details are presented in Appendix J). Overall, there is some evidence that incorporating physical activity within the academic curriculum or providing exercise breaks during the school day can have positive effects on student achievement and school engagement. Indeed, the evidence base was rated Well Supported for both publications.

Table 12: Summary of meta-analyses and systematic reviews (Physical activity)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Owen et al. 2016)	MA	38	K-12[a]	Integration of physical activity in academic lessons[b]	School engagement	0.40*, k=7 (RCTs)	0.27*, k=29 (children)	High	Well Supported
(Martin and Murtagh 2017)	SR	15	K-6[c]	Physically active teaching	Academic [d] Physical activity Health	Positive effects on learning outcomes in 4 of 4 studies (including 1 CRCT of 7-9 year olds[e])		High	Well Supported

[a] Ages 5-18 years; [b] includes analysis of specific strategies: providing physical exercise breaks in class, single bouts of exercise in 60 min prior to class; [c] Ages 5-12 years; [d] achievement defined as literacy, science, and mathematics; [e] Further details of this CRCT are provided in the next section, describing individual trials, *p<0.05

One meta-analysis (Owen et al. 2016) of 38 studies evaluated the extent to which physical activity is associated with school engagement. Overall, a small but positive association between physical activity and engagement was reported (ES=0.28) from studies of students in grades K to 12. A similar effect was observed for studies of 'children' (ES=0.27), with a larger effect for adolescents (ES=0.40). Data was not disaggregated for grades K to 3 students. A systematic review (Martin and Murtagh 2017) including 15 studies of classroom-based physical activity interventions *integrated* with academic content reported positive effects on learning, physical activity, and health outcomes. Although only four of the included studies focused on learning outcomes, positive effects on academic achievement were reported for all four studies. Importantly, one of these studies was a CRCT of children aged 7-9 years (Donnelly 2009); others were either not experimental or included older children. Overall, the evidence base for physical activity strategies to enhance educational outcomes was rated Well Supported because the meta-analysis and systematic review both identified positive effects on relevant outcomes, had low risk of bias, included a sufficient proportion of primary age students, and were based on studies conducted in countries considered similar to Australia. Table 13 describes the extent to which the effects observed in the meta-analysis and systematic review are likely to generalise to the early years of school.

Table 13: Effective physical activity strategies and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported		
• Providing physical exercise breaks during class (Owen et al. 2016)	Moderate	Possible
• Single-bouts of exercise in the 60 minutes before class (Owen et al. 2016)	Small to moderate	Possible
• Integration of physical activity with academic content (Martin and Murtagh 2017)	Not reported	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are reported. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

The proportion of primary age students was not reported *for specific strategies* included in the meta-analysis (Owen et al. 2016) and was very low for the systematic review (although it did identify one CRCT with children age 7-9 years). To supplement these findings a separate search was conducted for experimental and quasi-experimental trials of physical activity interventions implemented with children in the early years of school. Table 14 shows five relevant studies were identified (for details see Appendix J). One of these was the CRCT (Donnelly 2009) identified in the systematic review.

Table 14: Summary of CRCT and QES investigations (Physical activity)

RCTs and QESs								
Study	**Design**	**N**	**Grades**	**Strategy**	**Outcome Area**	**Results**	**Quality**	**Evidence Rank**
(Donnelly et al. 2009)	CRCT (24 schools)	1527	2-3	Physical Activity Across the Curriculum (PAAC) (90 min per week)	Academic (WIAT-II) BMI Physical activity	Positive effects on reading, math, spelling, and composite measures (all p<.01) ns p=0.001	Moderate	Promising
(Donnelly et al. 2017)	CRCT (17 schools)	584	2-3	A+ PAAC (55 min per week)	Academic Achievement (WIAT-II) Others [a]	All outcomes ns over a three-year period	Moderate	Not Supported
(Harvey et al. 2018)	QES (2 schools in larger CRCT)	66	2-3	A+ PAAC (100 min per week)	Learning behaviour (teacher-rated)	Significant improvement over time for intervention group: 0.43 (baseline to 3 months) 0.81 (3 to 6 months)	Moderate	Preliminary
(Li et al. 2014)	QRCT	83	2	Cup-stacking vs. traditional physical activity	Handwriting: • speed • quality	Positive trend: group x time interaction for writing speed (d=0.32, p=0.08)	Moderate	Unknown
(Mullender-Wijnsma et al. 2016)	CRCT (12 schools)	499	2-3	Fit & Vaardig op (20-30 minutes physical activity in maths and spelling 3x per week)	Mathematics speed General mathematics Reading Spelling	0.51* 0.42* ns 0.45*	Moderate	Promising

[a] BMI, waist circumference, cardiovascular fitness; ns= not significant; *p<0.05

Evidence from experimental and quasi-experimental evaluations

Five individual trials identified three different physical activity interventions. These included Physical Activity Across the Curriculum (PACC), Fit & Vardig Op, and cup-stacking programs. There is some (albeit limited) evidence that PAAC and Fit Vardig Op have positive effects on academic achievement and/or engagement. The evidence for these is described below.

Fit & Vardig Op

A CRCT of Fit & Vardig Op with grades 2 to 3 students at 12 schools in the Netherlands reported positive program effects on measures of mathematics speed and spelling (though not reading or overall mathematics scores) (Mullender-Wijnsma et al. 2016). This program involved the delivery of physically active mathematics and language lessons by qualified teachers. The main focus was on constant practice and repetition. For example, the children jumped on the spot 8 times to solve the multiplication sum "2 × 4." The physical exercises were aimed at moderate-to-vigorous intensity. It was delivered in 20-30 minute mathematics and spelling lessons, three times per week, over a period of 22 weeks for 2 years.

PAAC

The PAAC program involves physically active academic lessons of moderate-to-vigorous intensity. It has been evaluated in three trials with students in grades 2 and 3. Statistically significant positive effects were reported for two trials. A CRCT with 24 schools and more than 1500 students found intervention students scored significantly higher than controls (who received regular classroom instruction) on standardised measures of academic achievement in reading, mathematics, and spelling (Donnelly et al. 2009). In a quasi-experimental study including 2 schools, statistically significant improvements on teacher-rated learning behaviours were observed over time for intervention students (ES= 0.43 to 0.81, see Harvey 2018), but not comparison students. It is worth noting that in these trials, the PAAC program dose was relatively high (at least 90 minutes of moderate-to-vigorous physical activity). In contrast, a third trial tested the intervention delivered for only 55 minutes per week and found no statistically significant program effects (Donnelly et al. 2017).

Domain 4 Conclusions

The domain 'physical activity for academic achievement' was rated Supported.

Is there evidence that engaging students in physical activity can improve academic outcomes or school engagement?

Overall, strategies to incorporate physical activity in daily teaching practices have demonstrated positive effects for primary and secondary students. However, it is unclear which *specific strategies* are most effective for children in the early years. There is Promising evidence of positive effects from studies of programs that integrate moderate to vigorous physical activity with curriculum content (such as PAAC or Fit & Vaardig), but replication across multiple high quality RCTs is desirable.

Domain 5: Technology assisted teaching and learning

Digital technology is defined as the use of information and communication technologies to provide instruction in the classroom. This domain includes, for example, use of audio-visual material such as videos or animations to illustrate conceptual knowledge, interactive simulations or games that may be computer or web-based, computer-assisted learning programs that provide individual or group feedback and diagnostic or prescriptive instructional content, and computer-supported collaborative learning systems.

Research Questions:

5.1 Is there evidence that the use of digital technology in classroom instruction can improve student outcomes?

5.2 Which strategies for incorporating digital technology in classroom instruction have demonstrated positive effects on academic achievement in the early years of school?

Strength of the domain

Domain 5 was rated Well Supported. The evidence was strong and consistent with 10 meta-analyses identifying strategies rated Well Supported or Supported. Moreover, the evidence shows several strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were 12 meta-analyses/systematic reviews identified (see Table 15; details are presented in Appendix J). Overall, there is consistent evidence that the use of digital technology in the classroom can improve student outcomes. Indeed, the evidence base was rated Supported in eight of the publications and rated Well Supported for two. The most effective uses of technology and the extent to which the evidence is relevant to the early years of school is summarised in Table 16.

Table 15: Summary of meta-analyses and systematic reviews (Technology-assisted teaching)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Abrami, Borohkovski, and Lysenko 2015)	MA	9	PK-3	ABRACADABRA (interactive web-based balanced reading program)	Overall Phonics Phonological awareness Reading fluency Reading Comprehension Vocabulary Listening comprehension	0.17*, (#ES=73) 0.189*, (#ES=19) 0.324* (#ES=20) 0.078 (ns) (#ES=6) 0.065 (ns) #ES=6 0.108 (ns) , #ES=15 0.184 (ns), #ES=7		Moderate	Supported
(Cheung and Slavin 2012)	MA	84	K-12	Various educational technology[a]	Reading (innovative applications)	d'=0.16*, k=6	0.15, k=8 (kinder)	High	Supported
(Dietrichson et al. 2017)	MA	101 (9 CAI)	K-9	Computer assisted instruction	Reading & Mathematics (standardised composite)	0.11, k=9	NR	High	Unknown (this domain)
(Lee et al. 2013)	MA	58	K-12	Use of tech. in teaching[c]	Cognitive outcomes (NOS)	0.42, k=48	0.5 (K-3)	Moderate	Supported
(Ok et al. 2017)	SR	13 (7 rel.)	PK-12	Universal design for learning	*how* teachers use UDL-based tools and instruction is more important than the technology itself	NR	NR (38%pr, 23%EYS)	Moderate	Unknown
(Santangelo and Graham 2016)	MA	76 (4 rel.)	K-12	Technology use in handwriting	Handwriting legibility	0.85*, k=4	3 of the 4 studies were EYS	High	Supported
(Slavin et al. 2012)	SR	17	K-6	Science programs and practices (k=5 tech-based)	Science learning	d'=0.37, k=5	NR	Moderate	Supported

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Sokolowski, Li, and Willson 2015)	MA	24	1-8	Exploratory digitised environment	Mathematics	0.60*, k=24	0.61*,k=3 (lower elementary)	High	Well supported
(Sung, Chang, and Liu 2016)	MA	110	K-12+	Mobile devices [d]	Range of subjects [e]	0.52, #ES=419 0.65 (primary)	0.10, #ES=2 (Kinder)	Moderate	Supported
(Takacs, Swart, and Bus 2015)	MA	43	PK-6[f]	Tech-enhanced stories [g]	Story Comprehension Expressive vocabulary Receptive vocabulary Code-related skills Engagement	0.17*, #ES=38 0.20*, #ES=18 -0.08, ns, #ES=9 0.16, ns, #ES=14 0.26, ns, #ES=12	NR (%EYS NR)	High	Well supported
(Thomas et al. 2013)	MA	40	K-12+	Interactive CAI	Academic (NOS)	0.18*, #ES=55	0.19, ns, #ES=13 (primary) EYS: NR	Moderate	Supported
(Tingir et al. 2017)	MA	14	K-12	Mobile devices as part of curriculum	Achievement (science, math, reading combined)	0.55*, #ES=NR (primary)	NR (%EYS NR)	High	Supported

[a] examples include computers, multimedia, and interactive whiteboards; [b] ages 5-18 years, [c] examples include use of multimedia, PDAs, and integrated learning systems, [d] examples include PDAs, smart phones, and laptops; [e] Subjects included language arts, social studies, science and maths; [f] paper includes pre-school and/or elementary studies; [g] multimedia and inter-active features; d'= Procedures described by Lipsey and Wilson (2001) to estimate effect sizes when unadjusted standard deviations are not available; *p<0.05; rel.= number of studies relevant to this domain

A variety of digital technology strategies were identified. Well Supported uses of technology for instructional purposes included: exploratory digitised environments for mathematics (Sokolowski, Li, and Willson 2015), and technology-enhanced stories (Takacs, Swart, and Bus 2015). Supported strategies included: ABRACADABRA -an interactive web-based balanced reading program (Abrami, Borohkovski, and Lysenko 2015), the use of a variety of technologies such as computers, multimedia, interactive whiteboards (Lee et al. 2013; Cheung and Slavin 2012), use of mobile devices specifically, such as laptops, PDAs, and smart phones (Sung, Chang, and Liu 2016; Tingir et al. 2017), interactive Computer Assisted Technology (Thomas et al. 2013), hand-writing technology (Santangelo and Graham 2016), and technology-based science programs (Slavin et al. 2012).

Table 16 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in the identified meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 16: Effective digital technology strategies and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported		
Exploratory digitised environments for mathematics (Sokolowski, Li, and Willson 2015)	Moderate to Large	Very Plausible
Technology-enhanced stories (Takacs, Swart, and Bus 2015)	Small	Plausible
Supported		
Various technology for reading instruction (Cheung and Slavin 2012; Lee et al. 2013)	Very small	Very Plausible
ABRACADABRA (Abrami, Borohkovski, and Lysenko 2015)	Small to moderate	Very Plausible
Computer Assisted Instruction utilised in interactive rather than didactic learning environments (Thomas et al. 2013)	Small	Plausible
Technology graphically illustrating scientific concepts (Slavin et al. 2012)	Small to moderate	Plausible
Digitised hand-writing instruction (Santangelo and Graham 2016)	Large	Plausible
Use of mobile devices (Tingir et al. 2017; Sung, Chang, and Liu 2016)	Moderate	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary school specific results are reported. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Strategies rated Very Plausible

Well Supported and Supported technology strategies with positive effects considered very plausible to apply to the early years of school included: the ABRACADBRA program; exploratory digitised environments in mathematics instruction; and various technological tools such as computers, interactive whiteboards, and multimedia. The evidence is described next.

ABRACADABRA

ABRACADABRA (AbraA) is a supplementary interactive web-based software package that focuses on the development of reading skills. A meta-analysis of the intervention included nine studies of students in the early years (grades Pre-K to 3), with six experimental and three quasi-experimental studies (Abrami, Borohkovski, and Lysenko 2015). Results showed that compared with students receiving regular literacy instruction, AbraA demonstrated significant positive effects on a composite measure of early literacy skills (ES=0.17), and small-to-moderate effects for phonics (ES=0.19) and phonological awareness (ES=0.32) measures. Non-significant effects were observed for reading fluency, reading comprehension, vocabulary and listening comprehension. The evidence for use of AbraA was rated Supported rather than Well Supported because risk of bias was moderate. The findings were rated Very Plausible to apply to the early years of school because the meta-analysis was restricted to students in the early years (100%). However, it is worth noting that almost half of the included studies were quasi-experimental and utilised classrooms in which teachers were self-assigned to treatment conditions. Thus, it is possible differences in pedagogical capacities of intervention and comparison teachers account for some of the magnitude of effect.

Digitised mathematics instruction

One meta-analysis investigated digital mathematics instruction defined as an exploratory computerised environment where students work on a computer screen or iPod to formulate and mathematise patterns or solve problems (Sokolowski, Li, and Willson 2015). The meta-analysis included 24 studies of students in grades 1 to 8 including 10 experiments and 14 quasi-experimental designs. Overall, significant positive effects of moderate-to-large magnitude (ES=0.60) were reported. The use of digital mathematics instruction (compared with more traditional methods of instruction) also resulted in positive effects on mathematics outcomes for children in grades 1 to 3 (ES=0.61). The evidence for use of digital mathematics instruction was rated Well Supported because positive effects were observed, risk of bias was low, the proportion of studies with primary school students was high, and the countries in which studies were conducted were reported and comparable education systems. The findings were rated Very Plausible to apply to the early years of school because effect sizes specific to the early years of school were presented.

Various technological tools (computers, interactive whiteboards, multi-media)

Two meta-analyses investigated the use of "various technological tools"; defined as computers, interactive whiteboards and multimedia, personal digital assistants (PDAs) and integrated learning systems.

The first (Cheung and Slavin 2012) included 84 studies of students in grades K to 12 and comprised 25 experiments, 3 quasi-experimental designs and 47 matched-control designs. This meta-analysis focused on the use of computers, interactive whiteboards, and multimedia compared with traditional teaching methods. It reported small but significant effects for reading outcomes both in grades K to 12 (ES=0.16) and primary school (ES=0.10). The effect of using various technological tools was also investigated for children in kindergarten (ES=0.15). Although the effect for kindergarten children was not statistically significant it is worth noting the magnitude was similar to that for primary students.

Another meta-analysis (Lee et al. 2013) investigated the effect of various technological tools (namely the use of multimedia, PDAs, and integrated learning systems) compared with non-technological instruction on cognitive outcomes[17]. It comprised 58 studies of students in grades K to 12, including experiments and quasi-experimental designs. Overall, a positive effect of moderate magnitude was

[17]The types of cognitive outcomes included was not reported.

reported for students in grades K to 2 (ES=0.42). Additionally, statistically significant and positive effects were also reported for the use of various technological tools among children in grades K to 3 (ES=0.50).

For both meta-analyses, the evidence was rated Supported rather than Well Supported because there was moderate risk of bias, and the countries from which studies were drawn was not reported. However, findings were rated Very Plausible to apply to the early years of school because effects specific to studies of children in grades K to 3 were reported.

Strategies rated Plausible

Well Supported and Supported uses of technology with positive effects considered Plausible to apply to the early years of school included:

- Technology enhanced stories
- Computer Assisted Instruction (in interactive learning environments)
- Digitised handwriting instruction
- Technology-based science programs
- Mobile devices

Technology enhanced stories

There is evidence from one meta-analysis that the use of multimedia and interactive features in technology enhanced storybooks (compared with traditional storybook reading) can improve child literacy outcomes (Takacs, Swart, and Bus 2015). This meta-analysis included 43 studies of students in primary school including 42 experiments and 1 quasi-experimental design. It reported statistically significant positive (though small) effects for primary school students on measures of story comprehension (ES=0.17) and expressive vocabulary (ES=0.20). Non-significant results were observed for receptive vocabulary, code-related skills and engagement. The evidence for use of technology enhanced stories was rated Well Supported because the meta-analysis had very low risk of bias, focussed on primary age students, and reported positive findings from studies conducted in countries with comparable education systems. The findings were rated Plausible to apply to the early years of school because all studies were conducted in the primary years.

Computer-assisted instruction in an interactive learning environment

There is evidence from one meta-analysis of 40 RCTs and quasi-experimental studies that interventions using computer-assisted instruction (CAI) in interactive compared with didactic environments can improve student outcomes (Thomas et al. 2013). In this analysis, CAI was defined in terms of instructional methods that integrated computer technologies with explicit instructional objectives, rather than merely the presence of computer technology to support teaching and learning. Interactive learning environments were those in which the pedagogical approach promoted interaction between the learning environment, instructor and students. The approach is characterised by student choice with regard to pace and content of instruction, and the opportunity to receive elaborate feedback. Overall, the meta-analysis reported a small but significant effect on academic achievement for students from kinder to tertiary education (ES=0.18) and a similar (albeit non-significant) effect for primary school age participants specifically (ES=0.19). It is likely this result failed to reach significance because of lower statistical power. The evidence for CAI within an interactive learning environment was rated Supported rather than Well Supported because the meta-analysis did not report the countries from which the studies were drawn. Generalisability of findings to the early years of school was rated Plausible as results were presented separately for studies of primary school age students, but not for children in grades K to 3 specifically.

Use of technology in hand-writing instruction

A meta-analysis of 76 handwriting instruction studies identified four studies where use of technology was part of handwriting instruction (in three of these studies children were in grades K to 3)(Santangelo and Graham 2016). All four involved the use of a digitising tablet. In two studies, the tablet was used for copying letters and in the other two it was used as a part of comprehensive instruction. Three studies examined the effect of intervention on legibility and one on fluency. The mean weighted effect on writing legibility in these studies was large (ES=0.85). However, there are several reasons to be cautious about the large positive effect obtained from these studies. First, only one study compared the use of technology with usual instruction rather than no instruction. Second, two of the studies used technology as part of comprehensive instruction and the results could be attributed to aspects of the intervention other than technology. Third, the authors of the meta-analysis noted substantive risk of publication bias. The evidence for use of technology in handwriting instruction was rated Supported rather than Well Supported because the countries from which studies were drawn was not reported. Findings were rated Plausible to apply to the early years of school because 3 of the 4 studies contributing to the relevant effect targeted children in grades K to 3.

Use of technology for science instruction

A systematic review of 17 RCT or matched-control studies investigated a variety of programs and practices used in primary school science instruction. Of these, five studies examined technology programs (compared with either alternative programs or usual instruction)(Slavin et al. 2012). Importantly, the review excluded programs that were very brief or conducted in artificial circumstances (e.g. with unrealistic supports such as additional teaching staff), and studies of programs in which control groups did not study the same content as experimental groups. The five included programs utilised different forms of technology (e.g. animated videos, computer simulations, modelling, 3D virtual reality) but all utilised 'video or computer graphics to illustrate scientific processes, active inquiry using technology tools, integration of technology, teaching, and group work among students, and efforts to make science content motivating and relevant to students' (p. 20). Across the five studies the weighted mean effect size was small-to-moderate (ES=0.37). This result should be interpreted cautiously given that most of the studies relied on small sample sizes. Nevertheless, the evidence was rated Supported as the systematic review included appropriately designed studies of primary age students and had moderate risk of bias. Findings were rated Plausible to apply to the early years of school because although results are not presented separately for the early years, the proportion of relevant studies focusing on primary students was 100%.

Mobile Devices

There is evidence from two meta-analyses investigating the use of mobile devices (Sung, Chang, and Liu 2016; Tingir et al. 2017). One meta-analysis investigated the use of mobile devices (such as iPads, tablets, PDAs and smart phones) as part of the curriculum and compared with usual practice (Tingir et al. 2017). This meta-analysis included 14 studies of students in grades K to 12 including three experiments and 11 quasi-experimental designs. It reported a positive effect of moderate magnitude (ES=0.48) on a composite measure of achievement (comprising effects in science, mathematics and reading). A significant positive effect of moderate magnitude was also observed specifically among primary school students (ES=0.55).

Another meta-analysis investigated the effect of mobile devices (namely PDAs, smart phones, laptops) on students learning performance (Sung, Chang, and Liu 2016). This meta-analysis included 110 studies of students from kindergarten to graduate school. Combining results from all grades, moderate-to-large positive effects were observed across a range of subjects (ES=0.52), including language arts (ES=0.59), social studies (ES=0.78), and science (ES=0.58), with small-to-moderate effects in mathematics

(ES=0.34). Importantly, a positive effect of moderate-to-large magnitude was also observed for student learning performance in primary school (ES=0.65) and a small effect was also observed for children in kindergarten (ES=0.10). However, the meta-analysis did not report whether this result was statistically significant (or provide a 95% confidence interval), nor whether the result was from a single study including two outcome effects or two separate studies each contributing an outcome effect. It was also unclear which subject areas the contributing study(ies) had investigated.

Overall, the evidence for use of mobile devices was rated Supported because both analyses had moderate risk of bias and neither reported the countries in which the primary studies were conducted. The findings were rated Plausible to apply to the early years of school because results were presented at the primary school level in both, and although one included a kindergarten-specific effect it did not report whether the result was statistically significant, nor give an indication of the 95% confidence interval.

Domain 5 Conclusions

The domain 'Digital technology use in the classroom' was rated Well Supported.

Which types of digital technology use have demonstrated positive effects in the early years of school?

Specific uses of technology shown to be effective in the early years of school include:

- ABRACADABRA (a supplementary and interactive web-based reading skills program)
- Exploratory digitised environments in mathematics instruction
- Various technological tools such as computers, interactive whiteboards, and multi-media

All involve *interactive* instruction via technology rather than instruction via technology per se.

Domain 6: Physical environment design to optimise learning

The physical environment concerns spaces that can be altered by classroom teachers or school principals to improve student outcomes. Examples could include manipulating natural vs artificial lighting, access to green spaces, seating arrangements, location and/or accessibility of teaching materials, noise levels, and optimal heating and cooling. Safety, hygiene, and maintenance of buildings, grounds or equipment are not included in this domain as they are considered minimum standards that should be covered by legislation/regulation.

Research Questions:

6.1 Is there evidence that manipulating the physical environment is related to student academic outcomes?

6.2 If so, which aspects of the physical environment are related to student academic and cognitive outcomes? (i.e. correlational evidence)

6.3 If the physical environment does influence student achievement, which strategies can schools implement in the early years to ensure that the physical environment is most conducive to positive academic outcomes? (i.e. experimental evidence)

Strength of the domain

The evidence for Domain 6 was rated Preliminary. Only one (correlational) meta-analysis was identified. Similarly, a search for relevant experimental and quasi-experimental studies identified only three relevant publications. Of these, only one moderate quality RCT was identified. A summary of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

Table 17 summarises the identified meta-analysis (for details see Appendix J). It provides preliminary evidence that student outcomes are associated with the physical environment

Evidence from experimental and quasi-experimental evaluations

As the meta-analysis included only correlational evidence, a search for experimental and quasi-experimental trials was conducted. This search was confined to studies of children in the early years of school. Table 18 shows three studies were identified (see Appendix J for details).

Table 17: Summary of meta-analyses and systematic reviews (Physical environment design)

Study	Design	K Studies	Grades	Environment	Outcome Area	Overall Association	EYS Effect (grade/age)	Quality	Overall Evidence Rank

(Gunter and Sha	MA	18	K-12	Building condition & features	Science, language, mathematics composite	r=0.17*, b=0.06, #ES=92
						r=0.14*, b=0.10*, #ES=193
						r=0.14*, b=0.14*, #ES=92
						r=0.12*, b=0.10*, #ES = 594
							Primary:
							r=0.20*, b=0.11*, #ES=61
								Moderate	Supported

*p<0.05

The meta-analysis included 18 studies of K-12 students and reported a weak but statistically significant positive relationship between building condition and student achievement. A weak but positive association was also observed for primary students specifically. The correlational evidence was rated Supported rather than Well Supported because there was moderate risk of bias and the countries from which studies were drawn were not reported. Applicability of findings to the early years of schools was rated Plausible because a statistically significant effect size was reported for studies of primary age students.

Table 18: Summary of experimental studies (Physical environment design to optimise learning)

Experimental Studies								
Study	Design	N	Grades	Strategy	Outcome Area	Results	Quality	Overall Evidence Rank
(Amlani and Russo 2016)	REE	27	3	Acoustic panels & seat position	Word recognition Digit recall	No Panels vs Panels: 88% vs 79% recognition* 79% vs 69% recall* Panels ↓ performance* Seat proximity [a] (<2m vs >7m): 97% vs 68% recognition* 94% vs 52% recall*	Moderate	Unknown [b]
(Fisher, Godwin, and Seltman 2014)	REE	24	K	Classroom decoration	Time off-task Science learning	Negative effect: 0.85* Negative effect: 0.65*	Moderate	Unknown [b]
(Pfeiffer et al. 2008)	RCT	29	2	Disc-O-Sit for attention deficit	Behavioural regulation Meta-cognition (both teacher-rated)	η^2=0.23* η^2=0.32*	Moderate	Promising

[a] Seat proximity was assessed at four positions from less than 2 meters to more than 7 meters. Results are simplified here for brevity. [b] The evidence suggesting possible negative effects for acoustic panels was rated Unknown, because there were serious limitations to external validity. Likewise, the possible negative effect of classroom decoration was rated Unknown due to methodological limitations (i.e. confounding of condition with topic taught); REE= Repeated Exposure Experiment; *p<0.05

Practices rated Promising

Of the three strategies manipulating the physical classroom environment, only one was rated Promising. This was the use of a dynamic seating system for children with attentional difficulties.

Note: In domains where Supported and Well Supported strategies have been identified, only these have received plausibility ratings. However, given the relative dearth of good quality trials in this domain, a brief description of the sole promising strategy is provided.

Dynamic seating

In an unblinded randomised controlled trial, grade 2 students with attentional difficulties were assigned to use a dynamic seating system (the Disc 'o' Sit cushion) for two hours per school day for a period of two weeks (Pfeiffer et al. 2008). The seating system was a round air filled cushion designed to fit on a classroom chair and provide movement while seated. Teacher ratings of student behavioural regulation and meta-cognition demonstrated statistically significant and large effects favouring the treatment group, compared with the no treatment group (i.e. usual seating practices). However, caution interpreting results is warranted for several reasons including the possibility of biased teacher ratings and novelty effects. That is, it is possible the benefit may have resulted merely from the change in the environment rather than the cushion per se and may wear off when the cushion is no longer novel. Nevertheless, the practice is rated promising as it demonstrated a positive effect on a relevant outcome in a randomised controlled trial of moderate quality.

Domain 6 Conclusions

Overall, the domain 'Physical environment design to optimise learning' was rated Preliminary. Correlational evidence shows as there is an association between building conditions and features and student achievement.

Which types of classroom physical environment intervention have demonstrated positive effects in the early years of school?

No Well Supported or Supported classroom environment interventions demonstrating positive effects in early years of school were identified.

Further research is needed to more rigorously test various classroom environment interventions.

Domain 7: Class size and teacher-student ratios

Class size refers to the number of students in a classroom and instructor-student ratio refers to the number of instructors to students in a classroom.

Research Questions:

7.1 Is there evidence that class size or instructor-student ratios impact student academic outcomes?

7.2 If so, what is the optimal class size (or instructor-student ratio) for academic achievement in the early years of school?

Strength of the domain

Domain 7 was rated Supported as one moderate quality meta-analysis found consistently positive effects across several academic outcomes for class sizes of 22 or fewer students. Moreover, the meta-analysis showed the findings are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were two meta-analyses/systematic reviews identified[18] (see Table 19 for a summary and Appendix J for details). One of the meta-analyses demonstrated a positive effect between smaller class size and student academic achievement. No meta-analyses or systematic reviews examining optimal teacher-student ratio were identified. As such, a search for relevant experimental and quasi-experimental studies was conducted, however no publications met the inclusion criteria[19].

[18] A decision was made to include the seminal Glass & Smith (1979) meta-analysis because it has been very influential.
[19] A review of class size research was identified (Chingos, 2013) but not included because the review was not systematic. Similarly, several re-analyses of data from Project Star (a class size RCT conducted in Tennessee during the 1980s) were identified, but these did not meet inclusion criteria.

Table 19: Summary of meta-analyses (Class size and teacher-student ratios)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Glass and Smith 1979)	MA	77	NR	Smaller class size: 20 vs 40 students	Academic tests (standardised)	Glass Delta: 0.051	Primary: 10 percentile point advantage EYS: NR	Low	Unknown
(Shin and Chung 2009)	MA	17	K-12	Smaller classes (i.e. <22 students)	Overall Social Science Science Math Reading Writing	0.20*, #ES=120 0.20*; #ES= 9 0.15*, #ES=2 0.20*, #ES=34 0.19*, #ES=58 -0.09, #ES=1	0.19-0.24 (grade 1-3)	Moder ate	Supported

*p<0.05

Strategies rated Very Plausible

Class size of twenty-two or fewer students

Reducing class sizes to twenty-two or fewer students was examined in a meta-analysis of seventeen randomised and quasi-experimental studies (Shin and Chung 2009). Statistically significant positive effects of small magnitude were reported across standardised measures of reading, maths, and social science (0.19 to 0.20). Effect sizes were also reported separately for grades, with small but significant effects in grades 1 to 3 (ranging from 0.16 in grade 3 to 0.24 in grade 1). Additionally, the meta-analysis showed the overall effect was larger (0.20) in studies utilising random assignment, than those that were quasi-experimental (0.11). Overall, it suggests that students perform better in classes of twenty-two or fewer students than in larger classes. Unfortunately, it does not identify the optimum number of students per class. The strength of evidence was rated Supported rather than Well Supported because the quality assessment indicated moderate risk of bias. Applicability of findings to the early years of school was rated Very Plausible because effects were reported for children in grades 1 to 3.

Domain 7 Conclusions

The domain 'Class size and teacher-student ratios' was rated Supported.

Which class sizes and teacher-student ratios have demonstrated positive effects in the early years of school?

Class sizes of twenty-two or fewer students have demonstrated positive effects on standardised academic achievement measures.

PART II: THE SCHOOL CULTURE

This section focuses on the literature relevant to strategies that aim to improve the social climate or culture of the school environment. It includes approaches to (a) facilitating student empowerment, (b) ensuring the psychological well-being of students through social, emotional, and behavioural strategies, and (c) enhancing staff-student relationships. The three domains will be guided by specific research questions and specific strategies within those domains analysed and described.

Domain 8: Student empowerment and leadership

Student empowerment and leadership strategies are those that give students 'voice' or some influence in classroom or school-level decision-making. Examples of participation in decision-making may include for example: roles in councils, temporary school working groups, class decision-making, and school decision-making. Other types of empowerment such as autonomy or choice in learning materials (e.g. students choose books that interest them) or assessment items are not considered here. Rather, these are considered part of differentiated or tailored teaching strategies (see domain 2).

Research Questions:

8.1 Is there evidence that student empowerment or leadership strategies have positive effects on student social, emotional, and behavioural outcomes including school engagement?

8.2 Which student empowerment or leadership strategies have demonstrated positive effects on students in the early years of school?

Strength of the domain

Domain 8 was rated Unknown. No relevant meta-analyses or systematic reviews were identified. As such, a search for experimental and quasi-experimental studies of student empowerment or leadership strategies with children in the early years of school was conducted. No studies met inclusion criteria. A brief description of some non-experimental evidence is provided below.

Evidence from non-experimental studies

The extant literature relating to student empowerment and leadership in the primary school years appears limited to non-experimental study designs. For example, one review of single case reports, comparative case studies, and cross-sectional surveys suggests there is some evidence that student participation in school decision-making may have positive effects for a variety of outcomes including: life skills, self-esteem, social status, democratic skills, student-adult relationships and school ethos, academic achievement, facilities, rules or policies and health (e.g. Mager and Nowak 2012). Overall, there appears to be a substantive gap in the research literature pertaining to student empowerment and leadership strategies.

Domain 8 Conclusions

The domain 'Student empowerment and leadership' was rated Unknown.

Which student empowerment or leadership strategies have demonstrated positive effects in the early years of school?

No relevant strategies were identified in meta-analyses, systematic reviews, or individual experimental or quasi-experimental group comparison studies.

Domain 9: Social-emotional and behavioural strategies to promote a positive school climate

Social-emotional and behavioural strategies focus on the development of social awareness and relationship skills, emotion recognition and regulation, and appropriate and responsible behaviour. These may operate at the classroom or whole-school level.

Research Questions:

9.1 Is there evidence that school-based social-emotional and behavioural strategies can improve student outcomes?

9.2 Which social-emotional and behavioural strategies have demonstrated positive effects for children in the early years of school?

Strength of the domain

Domain 9 was rated Well Supported. The evidence was strong and consistent with seven meta-analyses identifying strategies rated Well Supported or Supported. Moreover, the evidence shows two strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were eleven meta-analyses/systematic reviews identified (see Table 20 for a summary and Appendix J for details). Strategies identified in two meta-analyses/systematic reviews were rated Well Supported, and rated Supported in five others. The remaining MA/SRs were rated Unknown because there was insufficient data related to children in the primary years of school.

The most effective social-emotional and behavioural strategies and the extent to which the evidence is relevant to the early years of school are summarised in Table 21.

Table 20: Summary of meta-analyses and systematic reviews (Social-emotional development)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Barbero et al. 2012)	SR	32	K-10[a]	Violence prevention & reduction programs	Behaviour Bullying Academic Social-emotional	NR[b]	Primary: 6 of 7 RCTS show positive effects EYS: NR	Moderate	Supported
(Carsley, Khoury, and Heath 2018)	MA	24	K-12	Various mindfulness interventions	Mental health & Wellbeing (composite)	0.24*@ post-test , #ES=21, 0.17 (ns) @ follow-up, #ES=6	0.22* @ post-test #ES=7 (6-10 years)	High	Supported
(Dietrichson et al. 2017)	MA	101 7 rel.	K-9	Psychological/ behavioural (k=7)	Composite	0.05 (ns), k=7	NR	High	Unknown (this domain)
(Durlak et al. 2011)	MA	213	K-12	Universal social & emotional learning programs	Overall Social-emotional Attitudes Prosocial Conduct Emotional Academic	0.30 0.57 0.23 0.24 0.22 0.24 0.27	Primary: NR EYS: NR (Grade did not moderate effects)	Moderate	Supported
(Korpershoek et al. 2016)	MA	47	PK-6	Classroom management	Overall Academic Behavioural Social emotional Motivational	0.22* 0.17* 0.24* 0.21* 0.08, ns	0.28* 0.23* 0.27* 0.25* 0.11, ns (PK-1)	High	Well Supported
(Maynard et al. 2012)	MA	28	K-12	Truancy programs	Attendance	0.46*, k=16	Primary: 0.16 (ns) in RCT/QES (k=2) EYS: NR	High	Unknown/ Not Supported (primary)
(Zenner, Herrnleben-Kurz, and Walach 2014)	MA	24	1-12	Mindfulness	Psychosocial [c]	0.40* (between-group studies, k=19)	2 relevant studies: 0.20 to 0.48, (ESs not pooled)	High	Promising
(Zoogman et al. 2015)	MA	20	K-12+[d]	Mindfulness		ES=del [f]	NR	High	Unknown

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect		EYS Effect (grade/age)	Quality	Evidence Rank
					Overall Objective [e] Non-objective Psychological symptoms Other symptoms Attention	0.227*, k=20 0.230*, k=6 0.255*, k=8 0.373*, k=15 0.207*, k=15) 0.280*, k=6)				
(Sklad, Diekstra, De Ritter, et al. 2012)	MA	75	K-12[g]	Social-emotional programs	- Academic Anti-social Mental health Self-image Pro-social Social-Emotional Substance use	Post 0.46*, k=10 0.43*, k=39 0.19*, k=13 0.46*, k=8 0.39*, k=6 0.70*, k=31 -0.09*, k=12	F/Up 0.26*, k=7 0.20*, k=16 -0.10*, k=11 0.07 , k=12* 0.12*, k=7 0.07*, k=15, -0.18*, k=24	Primary: ↑ social skills (d=0.67*) ↓antisocial behaviour (d=-0.59*) EYS: NR	Moderate	Supported
(Ttofi and Farrington 2011)	MA	44	K-12	Anti-bullying programs	Bullying Victimisation	ES=Odds ratio 1.36* 1.29*		Primary (≤10 yrs) 1.22, k=18 1.22, k=18	Moderate	Supported
(Whear et al. 2013)	SR	14	PK-6 [h]	Teacher-led universal SEB interventions	Overall SEB Academic Child-teacher relationships	Mixed [i]		Positive effects for specific trials (no effect sizes)	High	Well Supported

[a] Footnote authors state 5-16 years; [b] No effect size reported, but authors state most efficient strategies feature multi-disciplinary involvement, global focus, and target social and interpersonal skills, attitudes, and beliefs; [c] Outcome measures grouped as follows: Cognitive performance, emotional problems, stress and coping, resilience, "third person ratings" (In the domain of third person ratings, parent and teacher questionnaires were grouped, dealing with aspects such as aggressive or oppositional behavior, social skills, emotional competence, well-being, attention, and self-regulation); [d] Age 6-21 years; [e] Examples include: Objective measures (psychophysiological measures, attention and behavioural tasks), Non-objective measures (teacher-,parent-,or child-report); [f] Del is a measure of the difference in pre-post ES between groups, in this case comparison between mindfulness strategies and alternative treatments. Use of del compares the change over time in the two groups which is typically the outcome of interest; [g] 'primary and secondary', grades NR; [h] Ages 3 to 12 years; [i] Statistically significant positive effects for 20 outcomes, significant effects for 8 of 12 studies with behavioural outcomes (7 in a positive direction),few effects for social, emotional and academic outcomes ; *p<0.05

A variety of school-based strategies targeting child social-emotional and behavioural outcomes were identified. Well Supported strategies were classroom management interventions (Korpershoek et al. 2016; Whear et al. 2013). Supported strategies were violence prevention and reduction programs (Ttofi and Farrington 2011; Barbero et al. 2012), mindfulness interventions (Carsley, Khoury, and Heath 2018; Zenner, Herrnleben-Kurz, and Walach 2014), and universal social-emotional skills programs (Sklad, Diekstra, De Ritter, et al. 2012).

Table 21 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 21: Effective social-emotional development strategies and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported		
Teacher training in classroom management programs and strategies (Whear et al. 2013; Korpershoek et al. 2016)	Small	Very Plausible
Supported		
Anti-bullying, violence prevention and reduction programs (Barbero et al. 2012; Ttofi and Farrington 2011)	Unclear	Plausible
Mindfulness interventions (Carsley, Khoury, and Heath 2018; Zenner, Herrnleben-Kurz, and Walach 2014)	Small to moderate	Very Plausible
Specific/Universal SEB programs (Sklad, Diekstra, De Ritter, et al. 2012; Durlak et al. 2011)	Small to moderate	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are presented. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Strategies rated Very Plausible

Well Supported and Supported social-emotional and behavioural interventions with positive effects considered Very Plausible to apply to the early years of school included: teacher training in classroom management and mindfulness interventions.

Teacher training in classroom management

Definitions of effective classroom management emphasise the importance of various strategies teachers use to create an environment conducive to instruction. Examples include promotion of appropriate social behaviour, engagement in academic tasks, and implementation of strategies to assist children with behavioural problems (Korpershoek et al. 2016).

Evidence supporting various classroom management strategies was identified in one meta-analysis (Korpershoek et al. 2016) and one systematic review (Whear et al. 2013). The meta-analysis included 47 randomised and quasi-experimental studies investigating universal classroom management programs and strategies for primary school age children (Korpershoek et al. 2016). Importantly, small

but significant positive effects were demonstrated for academic (0.23), behavioural (0.27) and social-emotional outcomes (0.25) for children in grades pre-K to 1.

In addition to these general findings, the meta-analysis (Korpershoek et al. 2016) compared various strategies and specific programs. With regard to strategies, results indicated that programs with a strong social-emotional development component (e.g. empathy and social skills acquisition), had larger overall effects than those without (ES= .24 vs .15, p=.05 across all outcomes; 0.25 vs 0.04, p<.01 for social-emotional outcomes). Programs with a focus on teacher actions (e.g. use of rules, disciplinary strategies), or student behaviours (e.g. group contingencies or self-regulation skills), also reported significant effects (ES = .20, .21, respectively) though these were not larger than programs without such a focus. With regard to specific programs, the analysis compared results for School-Wide Positive Behaviour Support (SWPBS, k=3), Promoting Alternative Thinking Strategies PATHS (k=10), the Good Behaviour Game (GBG, k=4), Second Step (k=3) and Zippy's Friends (k=3). Overall, all programs except SWPBS demonstrated small-to-moderate effects (0.19 to 0.29), and the PATHS program had the strongest evidence base with larger effects and more studies.

The systematic review included 14 studies, also utilising RCT and quasi-experimental study designs with primary school children (Whear et al. 2013). It found several programs had positive effects on behavioural outcomes. More specifically, of the 12 studies that included behavioural outcomes, eight reported significant effects; seven of these were in a positive direction, with only one indicating a negative outcome. Specific named programs with positive effects on behavioural outcomes included the Incredible Years Teacher Classroom Management program; Good Behaviour Game, and the Proactive Classroom Management Program. Some classroom management programs have also demonstrated positive effects on academic achievement (e.g. The Proactive Classroom Management Program) and prosocial behaviour (e.g. The Incredible Years Teacher Classroom Management program). Results from the systematic review identified several evaluations conducted with children in the early years of school and two of these reported statistically significant positive program effects in RCTs. One reported a positive effect on prosocial behaviour for the Incredible Years Teacher Classroom Management program (Shernoff and Kratochwill 2007), and the other reported a positive effect for on-task behaviour with the Good Behaviour Game (Leflot et al. 2010). The extent to which findings for teacher training in classroom management are likely to generalise to the early years of school was rated Very Plausible because both the meta-analysis and systematic review reported positive effects specific to children in the early years of school.

Mindfulness strategies

Mindfulness strategies teach participants to continually and non-judgmentally focus their attention on present moment experience, noticing current thoughts, emotions, or body sensations. Mindfulness strategies included in the meta-analyses described here included structured activities to focus attention on present-moment physical and mental activity.

Three meta-analyses (Carsley, Khoury, and Heath 2018; Zoogman et al. 2015; Zenner, Herrnleben-Kurz, and Walach 2014) investigated the use of mindfulness interventions for school children. The most recent comprised twenty-four studies including children in grades K to 12 (Carsley, Khoury, and Heath 2018). Small but statistically significant positive effects emerged for a variety[20] of school-based mindfulness interventions (compared with active controls, usual practice, and waitlist comparison groups) on composite measures of mental health and wellbeing. Effects were also investigated

[20] School-based mindfulness intervention studies included Mindfulness-Based Stress Reduction (k=3), Mindfulness-Based Cognitive Therapy for Children (k=3), single-component Mindfulness-Based Stress Reduction (k=5), and 'other' types of mindfulness intervention (k=9).

specifically for children aged 6-10 years, with results also showing mindfulness interventions resulted in small but positive effects at post-test assessments (ES=0.22).

A second meta-analysis (Zenner, Herrnleben-Kurz, and Walach 2014) investigated school-based mindfulness interventions for children in grades 1 to 12. It included 24 studies (18 of which were RCTs or quasi-experimental studies). Analysis of controlled between group studies found large positive effects on measures of cognitive performance (ES=0.80), small-to-moderate effects for measures of stress (ES=0.39) and resilience (ES=0.36), and small effects on measures of emotional problems (ES=0.19). Although less than half of the studies included in the analysis targeted exclusively primary age children, the study did identify two RCT studies of children in the early years of school. One study of 64 students in grades 2 and 3 reported positive effects of a mindfulness intervention on both parent (ES=0.39) and teacher (ES=0.20) composite ratings of behaviour and executive function. The other study included 194 students in grade 1 -3 and reported small-to-moderate effects on measures of attention (ES=0.60) and emotional problems (ES=0.39).

A third meta-analysis (Zoogman et al. 2015) of mindfulness interventions for children under 18 years of age included 20 studies, most using non-clinical samples. Small and small-to-moderate effects of mindfulness training were demonstrated across a range of outcomes including objective and non-objective measures (e.g. teacher, parent, or child-reported outcomes), psychological symptoms (e.g. anxiety, aggression) and others (e.g. attention). It included 13 RCTs and one quasi-experimental study. Although the proportion of studies covering children in primary school grades was not reported, statistical analysis indicated that age did not moderate effects.

Overall, the evidence for use of mindfulness strategies was rated Supported as all three meta-analyses were high quality and reported positive effects. The overall rating was not Well Supported because the proportion of primary school age students was low in one meta-analysis (Zenner, Herrnleben-Kurz, and Walach 2014) and not reported in another – although age was investigated as a possible (non-significant) moderator (Zoogman et al. 2015). However, findings were rated Very Plausible to apply to the early years of school because positive results specific to children in the early years of school were identifiable in two of the three meta-analyses, though only pooled in one meta-analysis.

Strategies rated Plausible

Violence prevention and reduction programs

One meta-analysis (Ttofi and Farrington 2011) and one systematic review (Barbero et al. 2012) investigated the use of anti-bullying and anti-violence programs. The meta-analysis included forty-four programs implemented with students from kinder to high school. Small positive effects of anti-bullying programs (compared with no program) were observed on both bullying and victimisation measures. Importantly, these occurred in studies of children 10 years or younger (odds ratios were 1.22 on both measures for this age group).

A systematic review of anti-violence programs (Barbero et al. 2012) that directly target students (rather than teachers or parents) also reported positive program effects for students. This review included 32 publications, of which two were meta-analyses of RCTs, one was a systematic review, and 12 were original research studies utilising RCT designs. Of the two meta-analyses of RCTs, one focused on children aged 5-12 years, and identified as aggressive or at-risk of being so. It reported significant reductions of small-to-moderate magnitude (ES=0.41) in aggressive behaviour for 34 of 56 RCTs, with results maintained in seven studies that included follow up measurements. The results from the nine RCTs evaluated seven programs in primary schools. Of these, six reported positive and statistically

significant effects on outcomes such as externalising, anti-social behaviours, peer friendliness, social anxiety, self-esteem, self-efficacy, bystander behaviour, physical and verbal bullying, as well as risk taking behaviours (drug use and sexual activity) at 4 year follow up. Results from only one of the included trials was specific to the early years (participants were 8 years of age). Overall, the authors state that the most effective strategies appear to be those that target social and interpersonal skills and modify attitudes and beliefs. They recommend that programs (a) involve parents and all professional disciplines within the school, (b) are adapted to the social and cultural characteristics of the school, and (c) incorporate continuity strategies such as reminder sessions or integration of programs with the academic curriculum.

The evidence for use of anti-bullying and anti-violence programs was rated Supported rather than Well Supported because both the meta-analysis and systematic review had moderate risk of bias. The findings were rated Plausible to apply to the early years of school because both publications presented results for studies of primary school age children, but did not pool results across studies of children in the early years of school specifically.

Universal school-based social-emotional learning programs

Universal school-based programs are those that are delivered to the general school population, rather than targeting children with specific risk factors (such as economic disadvantage) or established difficulties. One meta-analysis (Sklad, Diekstra, De Ritter, et al. 2012) investigated the use of universal school-based programs in which primary or secondary students were taught at least one social-emotional skill. This analysis included 75 studies (42 experimental and 33 quasi-experimental) published between 1995 and 2008. Overall, positive program effects emerged across a range of outcomes including academic achievement, anti-social behaviour, mental health, self-image, pro-social behaviour, social-emotional well-being, and substance use. Effect sizes ranged from small (e.g. mental health, ES=0.19) to moderate-large (social emotional well-being, ES=0.70) at immediate post-test, and were generally maintained (albeit smaller in magnitude) at follow up tests administered at least 6 months after program completion. The analysis also reported substantive post-test effects for studies of primary school students, with moderate effects indicating superior social skills (ES=0.67) and less anti-social behaviour (ES=-0.59) among intervention compared with control group children.

A second meta-analysis (Durlak 2011) investigating school-based social-emotional development programs implemented with children from grades K-12 included 213 studies. All were experimental (47%) or quasi-experimental (57%) and published between 1970 and 2007. Overall, social-emotional development programs had small to moderate positive effects on a range of immediate outcomes (0.30) including social-emotional skills (0.57), attitudes toward self and others (0.23), prosocial behaviour (0.24), conduct problems (0.22), emotional distress (0.24), and academic achievement (0.27). Significant and positive (albeit smaller) effects were also maintained in follow up assessments, at least 6 months from implementations, and on average conducted 12 months post-intervention. Programs delivered by teachers had significant and positive effects of similar magnitude to the overall findings. Programs that were characterised by sequenced, active, focused, and explicit instruction (SAFE[21]) had statistically significant positive effects on all outcomes. In contrast, programs not characterised by these features had significant positive effects on only three of six outcomes, and were smaller in magnitude. Although the meta-analysis did not present results separately for primary school children, and the proportion of studies focusing exclusively on primary students was 56%, further

[21] Sequenced = program uses connected and co-ordinated sets of activities; Active = program uses active forms of learning; Focused = program includes at least one component devoted to developing personal or social skills; Explicit = program targets specific social-emotional skills rather than positive development in general

analysis indicated that student grade did not moderate effects, though student age was significantly and negatively related to measures of social-emotional skills (r=-0.27), but not other outcomes. Thus, it appears that, overall, the strategies implemented with primary school students were as effective as those implemented with older students across most outcomes.

The evidence for use of programs specifically targeting social-emotional skills was rated Supported rather than Well Supported because the meta-analyses had moderate risk of bias. Findings were rated Plausible to apply to the early years of school because positive effects specific to primary school children were reported in one meta-analysis, and grade was shown not to moderate effects in the other.

Domain 9 Conclusions

The domain 'Social-emotional and behavioural strategies to support a positive school climate' was rated Well Supported.

Which social-emotional and behavioural strategies have demonstrated positive effects for children in the early years of school?

Social-emotional and behavioural strategies shown to be effective in the early years of school include:

- Teacher training in classroom management (such as the Incredible Years Teacher Classroom Management Program, and Good Behaviour Game).
- Mindfulness programs.

Domain 10: Teacher student relationships

This domain concerns the importance of teacher-student relationships to student outcomes and potential strategies to ensure that these relationships are mutually respectful. Included studies may measure teacher-student relationship quality by (a) either teacher or student perception or (b) direct observation of teacher-student interactions.

Research Questions:

10.1 Is there evidence that teacher-student relationships are associated with student outcomes? (academic and non-academic)

10.2 Which types of strategies to improve teacher-student relationships have demonstrated positive effects in the early years of school?

Strength of the domain

The evidence for Domain 10 was rated Preliminary. Correlational evidence from three meta-analyses/systematic reviews shows teacher-student relationships are associated with student outcomes, but experimental evidence was severely limited. Of four teacher-student relationship interventions identified as applicable to the early years of school, only two demonstrated positive main effects. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were five meta-analyses/systematic reviews identified (see Table 22 for a summary and Appendix J for details). Among these, there is consistent evidence that teacher-student relationships are related to student outcomes. Indeed, the correlational evidence base was rated Supported or Well Supported for three of the publications, and Unknown for the remaining two.

However, the evidence base from meta-analyses for the types of strategies that improve teacher-student relationships is notably sparse. Indeed, only one meta-analysis reported an effect size for strategies involving a teacher-student relationship component. Unfortunately, only two studies contributed to the reported effect size (which was non-significant and small). As such, a search for relevant experimental and quasi-experimental studies was also conducted. Table 24 shows four relevant publications were identified. Strategies tested with students in the early years of school and rated Promising are then described.

Table 22: Summary of meta-analyses (Teacher-student relationships)

	Design	K Studies	Grades	Relationship or Strategy	Outcome Area	Overall Association	EYS Effect (grade/age)	Quality	Evidence Rank
(Cornelius-White 2007)	MA	119	PK-12+	Association: Teacher-student relationship	Cognitive, Affective, Behavioural (composite)	r=0.36*	NR	Moderate	Unknown (RQ1)
(Korpershoek et al. 2016)	MA	54	1-6	Strategy: Classroom management	Academic, Behavioural, Social Emotional	d=0.13, ns, k=2 with teacher-student r/ship component	NR (but all primary)	High	Unknown (this domain)
(Lei, Cui, and Chiu 2016)	MA	57	K-6	Association: Affective teacher-student relationship	Externalising problems	+ve r/ships r=-.26* -ve r/ships r=0.55	Age 6-9yrs: +ve r/ship r= -0.28* -ve r/ship r=+.56* Kinder: +ve r/ship r=-0.19* -ve r/ship r=0.48*	Moderate	Supported (RQ1)
(Roorda et al. 2017)	MA	189	PK-12	Association: Affective teacher-student r/ship	Student Achievement [a]	0.14* +ve r/ships -0.12* -ve r/ships (total effects)[b]	Primary (k=105): β = 0.07* +ve r/ship β = -0.07*-ve r/ship	Moderate	Supported (RQ1)
(Vandenbroucke et al. 2018)	MA	23	PK-6 [c]	Association: teacher-student interaction [d]	Overall Executive function Working memory Inhibition Cognitive flexibility	R=0.09* (k=23) r=0.11* (k=3) r=0.10* (k=7) r=0.08* (k=17) r=0.00, ns (k=3)	NR (but sample 96% EYS)	High	Well Supported (RQ1)

[a] authors did not list or distinguish between different subjects; [b] These are effects for the Total model (i.e. summing direct path from relationship to achievement and indirect path from relationship to achievement via student engagement; [c] ages 2 to 12 years; [d] e.g., closeness, conflict, classroom organisation *p<0.05, RQ1=Research Question 1

Several meta-analyses identified significant associations between teacher-student relationships and student outcomes. Well Supported associations included teacher-student interaction quality and student executive function (Vandenbroucke et al. 2018). Supported associations included the affective nature of the teacher student relationship with student externalising symptoms (Lei, Cui, and Chiu 2016), and with student academic achievement (Roorda et al. 2017).

Table 23 summarises associations rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for associations rated Very Plausible and Plausible is summarised below.

Table 23: Teacher-student relationship association with student outcomes and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported Relationships		
• Teacher-student interaction and student cognitive functioning (Vandenbroucke et al. 2018)	Small	Very Plausible
Supported Relationships		
• Teacher-student relationship and student externalising problems (Lei, Cui, and Chiu 2016)	Small to moderate	Very Plausible
• Teacher-student relationship and student academic achievement (Roorda et al. 2017)	Small	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on students the in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are presented. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Associations rated Very Plausible

Cognitive Functioning

The association between teacher-student interactions and student cognitive functioning was investigated in one meta-analysis (Vandenbroucke et al. 2018). It included 23 studies of primary school students and reported several small but statistically significant associations (all *r*s <0.12). Specifically, positive associations were reported for teacher-student interactions with measures of executive function (ES=0.11), working memory (ES=0.10), and inhibition (ES=0.08). Although effect sizes specific to the early years of school were not presented, the proportion of studies with children exclusively within the target age range was very high (96%). For this reason, the generalisability of findings to the early years of school was rated Very Plausible.

Externalising Problems

The association between teacher-student interactions and externalising problems was investigated in one meta-analysis (Lei, Cui, and Chiu 2016). It included 57 studies of primary school students and reported a statistically significant but small negative association between positive teacher-student relationship and externalising problems (r=-0.26). Furthermore, effect sizes specific to the early years of school were presented. For children in kindergarten there was a small negative association between positive teacher-student relationship and externalising problems (r=-0.19) and a moderate positive

association between negative teacher-student relationship and externalising problems (r=0.48). Similar results were found when the analysis was restricted to 6-9-year-old students (see Table 22).

Associations rated Plausible

Academic Achievement

The association between teacher-student interactions and student achievement was investigated in one meta-analysis (Roorda et al. 2017). It included 189 studies of primary and high school students and reported small but statistically significant associations. A small positive association was reported for positive teacher-student relationship with student achievement (r=0.14). A small negative association was reported for negative teacher-student relationship with student achievement (r= -0.12). Although effect sizes specific to the early years of school were not presented, results specific to the primary years were (k=105). Consistent with results for the overall sample, there were small but significant associations in the primary years such that positive teacher-student relationships were associated with stronger student achievement (r=0.07) and negative teacher-student relationship with poorer student achievement in (r= -0.07). As the proportion of studies with children exclusively within the target age range was not recorded, the generalisability of findings to the early years was rated Plausible.

Table 24: Summary of RCT and CRCTs (Teacher-student relationships)

RCTs								
Study	**Design**	**N**	**Grades**	**Strategy**	**Outcome Area**	**Results**	**Quality**	**Evidence Rank**
(Abry et al. 2013)	CRCT (24 schools)	239 Teachers (I:132 C: 107) S:NR	3-4	Responsive Classrooms (teacher PD)	Teacher-student interaction quality (Obs)	Direct effect: β= -0.68* Indirect (via implementation fidelity): β =0.52*	High	Unknown
(Cappella et al. 2012)	RCT	36 Teachers 347 Students (I: 169, C:178)	K-5	BRIDGE program (teacher PD: interacting with behaviourally challenging students)	Teacher-student relationship closeness (TR) Others[a]	ES=0.47*	High	Promising
(Fernandez et al. 2015)	RCT	12 Teachers 118 Students (I: 64, C: 54)	K-1	Teacher Child Interaction Training	Teacher skills [b] (Obs) Teacher distress (Obs) Student behaviour (TR)	Post-test: 0.48 to 1.34 0.90* 0.30 (ns) Follow-up 1.08 to 1.50 0.62* 0.22 (ns)	Moderate	Promising
(Spilt et al. 2012)	RCT	16 Teachers 32 Students	K	Relationship Focused Reflection Program	Closeness (TR) Conflict (TR) Teacher behaviour (0bs) Child behaviour (Obs)	No significant Tx effects Subgroup: +ve Tx effects for high efficacy teachers (on Obs sensitivity and TR conflict)	Moderate	Not Supported

[a] academic self-concept, and peer victimisation; [b] labelled praise, behaviour descriptions, reflections; C=Comparison group; I=Intervention group;, S=Students, T= teachers, TR=teacher rated, Tx= treatment; Obs=Observed; *p<0.05

Strategies rated Promising

Note: In domains where Supported and Well Supported strategies have been identified, only these have received plausibility ratings. However, given the relative dearth of good quality trials in this domain, an applicability rating and brief description of Promising strategies is provided.

Promising strategies rated Very Plausible

Teacher Child Interaction Training

Teacher Child Interaction Training (TCIT) is a classroom-based program, designed to provide teachers with behaviour management skills that foster positive teacher-student relationships. It consists of two phases. The first is child-directed interaction (CDI) which is focused on teaching skills thought to strengthen the teacher student relationship and increase student prosocial behaviour. The second phase is teacher-directed instruction (TDI) which is focused on teaching specific, limit-setting techniques to decrease problem behaviours. Training duration is approximately 11 weeks. The TCIT program was tested in a RCT (Fernandez et al. 2015) including eleven teachers of 118 students in grades K to 1. Moderate-to-large effects favouring intervention teachers compared with no intervention, were reported on observer ratings of teacher skill (ES=0.48 to 1.34) and teacher distress (ES=0.90) at post-test. Similarly, at 1-month follow-up moderate-to-large effects favouring intervention teachers were reported on observer ratings of teacher skill (ES=1.08 to 1.50) and teacher distress (ES=0.62). There was also some indication that the strategy had positive effects on teacher-rated student behaviours but the effects were not statistically significant. The evidence for this strategy was rated Promising because positive effects were observed in a moderate quality RCT. However, results should be interpreted cautiously as there were several serious design limitations. For example, the teacher sample was very small and measures of student behaviour were teacher-reported (rather than blind observation).

Promising strategies rated Plausible

BRIDGE Program

BRIDGE is a teacher consultation and coaching program which aims to increase effective classroom interactions with behaviourally challenging students. The BRIDGE program was compared with a reflective teaching control group in a RCT (Cappella et al. 2012) including 36 teachers, and 347 students in grades K to 5, in 5 schools. Teachers received on average 4.47 observation and coaching sessions, and 3.50 consultation meetings (25-30 minutes each). A statistically significant effect of moderate magnitude indicated a positive effect on teacher-reported teacher-student relationships at post-test (ES=0.47). Statistically significant effects of small-to-moderate magnitude also indicated positive effects on academic self-concept (ES=0.31) and peer victimisation (ES=0.31). There was no significant effect found for aggressive student behaviour. The evidence for this strategy was rated Promising because positive effects were observed in a high-quality RCT. While it is plausible that results would hold for teachers of students in the early years of school specifically, results should be interpreted with some caution given that the proportion of children exclusively within the target age range was not reported. It is possible that the strategies teachers implemented work better with older (or younger) students. Several design limitations should also be considered. For example, there is the potential for selection bias given that the number of teacher and target child participants was a relatively low proportion of the available population. Another limitation is that the study did not assess student ratings of the teacher-student relationship or direct measures of student outcomes (e.g. academic achievement).

Domain 10 Conclusions

The domain 'teacher-student relationships' was rated Preliminary.

Correlational evidence shows there is an association between teacher-student relationships and various student outcomes.

Which strategies targeting teacher-student relationships have demonstrated positive effects in the early years of school?

No Well Supported or Supported teacher-student relationship strategies demonstrating positive effects in early years of school were identified.

Further research is needed to test the replicability of promising effects shown in preliminary trials of teacher-student relationship strategies such as Teacher Child Interaction Training and BRIDGE

PART III: PROVIDERS AND PARTNERSHIPS

This section focuses on the literature relevant to interventions or strategies that aim to develop the skills of school staff and form collaborative partnerships with parents and local communities. In each domain, the types of interventions or strategies considered relevant are defined, and research questions to guide the review are provided. Then results from the literature search are presented and summarised.

Domain 11: Staff and leadership development

Professional development (PD) covers 'a wide variety of specialised training, formal education, or advanced professional learning intended to help administrators, teachers, and other educators improve their professional knowledge, competence, skill, and effectiveness' (see https://www.edglossary.org/). We use the term 'staff' to refer to teachers and teacher aides, and 'leadership' to refer to principals, leading teachers, or administrators.

Research Questions:

11.1 Is there evidence that teacher or school leadership PD strategies can have positive effects on student outcomes?

11.2 Which teacher and school leadership PD strategies have demonstrated positive effects for students in the early years of school?

Strength of the domain

Domain 11 was rated Well Supported. There is consistent evidence that teacher professional development can improve both academic and social-emotional student outcomes. Five publications identified evidence rated Supported or Well Supported. Moreover, positive effects of teacher PD were demonstrated for children in the early years of school. In contrast, a notable dearth of strong research evidence pertaining to school leadership development exists. Indeed, only one relevant meta-analysis and one RCT were identified. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were seven meta-analyses/systematic reviews identified (see Table 25 for a summary and Appendix J for details). Six of these pertained to teacher professional development, and one to leadership development. While the evidence for strategies identified in five of the six teacher PD publications was rated Well Supported or Supported, the evidence for leadership strategies was confined to one publication and rated Unknown. The most effective teacher PD strategies and the extent to which the evidence is relevant to the early years of school are summarised in Table 26.

Table 25: Summary of meta-analyses and systematic reviews (Professional development)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Blank and De Las Alas 2009)	MA	16	K-12	In-service Mathematics & Science PD	Mathematics Science	0.13* to 0.21*#ES=68 0.05 to 0.18, ns Primary math: 0.27* to 0.32* Primary science: ns	EYS: NR	High	Well Supported
(Dietrichson et al. 2017)	MA	101 18 rel.	K-9	Various: coaching/mentoring (k=10), professional development (k=8)	Academic (reading/ math)	0.16* , k=10 (coaching) 0.07, ns, k=8 (PD)	NR	Moderate	Unknown (this domain)
(Dunst, Bruder, and Hamby 2015)	SR	15	PK-12	In-service PD [b]	Academic Behavioural Teacher outcomes [c]	Core components identified [d]	NR	Moderate	Supported
(Kraft, Blazar, and Hogan 2018)	MA	60	PK-12	Teacher coaching [e]	Student achievement Teacher instruction	0.18, k=31 0.49*, k=43	0.11* (PK)	Moderate	Supported
(Leithwood and Sun 2012)	MA	79	K-12	Transformational school leadership	Reading Mathematics	r=0.15* r=0.18*	NR	Low	Unknown
(Slavin et al. 2012)	SR	17 8 rel.	K-6	Programs and practices in elementary science: Sub: Inquiry-based programs emphasising PD (k=8)	Science	d'=0.30, k=8	NR	Moderate	Supported
(Whear et al. 2013)	SR	14	PK-3[g]	Teacher training in classroom management programs	Academic Social, Emotional, Behavioural Teacher effects	Statistically significant positive effects for 20 outcomes. Evidence strongest for behavioural outcomes. Less evidence for social, emotional, and academic outcomes (PK-3)		High	Well Supported

[a] 5 to 18 years; [b] Focus of the training was on promotion of (a) different types of instructional or behavioural practices, (b) teacher understanding and use of content knowledge or skills and (c) practices to support teachers' confidence in their teaching practices; [c] teacher attitudes & beliefs, content knowledge, instructional or behavioural practices [d] Sufficient duration and intensity, extended follow-up supports and opportunities to reinforce content knowledge or practice; [e] examined content-specific, reading-specific, and general coaching [f] Likely about principals [g] ages 2 to 9 years; d'= Procedures described by Lipsey and Wilson (2001) to estimate effect sizes when unadjusted standard deviations were not available;*p<0.0

Several effective teacher professional development strategies were identified. Well Supported strategies included teacher training in classroom management (Whear 2013) and in-service PD for teaching maths and science (Blank 2009). Supported strategies included in-service teacher coaching (Kraft 2018; Dunst 2015) and use of inquiry-based science programs emphasising PD such as coaching and pedagogical content knowledge (Slavin 2012).

Table 26 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 26: Effective professional development strategies and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported		
In-service mathematics and science professional development (Blank and De Las Alas 2009)	Small	Plausible
Teacher-training in class-room management (Whear et al. 2013)	Not reported	Very Plausible
Supported		
Teacher coaching and mentoring (Dunst, Bruder, and Hamby 2015; Kraft, Blazar, and Hogan 2018)	Small	Very Plausible (Kraft, Blazar, and Hogan 2018)
Inquiry-based science programs with PD emphasised (Slavin et al. 2012)	Small to moderate	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are presented. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Teacher Training in Classroom Management

A systematic review of teacher training in classroom management (Whear et al. 2013) identified several positive effects of this type of professional development. It included 14 studies utilising RCT and quasi-experimental study designs. Although statistically significant treatment effects were observed for only 21 outcomes (of a possible 151), positive effects were consistently identified on student behavioural outcomes. Indeed, of the 12 studies including behavioural outcomes, eight reported significant effects and seven of these were in a positive direction. Specific programs with positive effects on behavioural outcomes were the Incredible Years Teacher Classroom Management program (IYTCM); Good Behaviour Game (GBG), and the Proactive Classroom Management Program (PCMP). The PCMP also had a positive effect (of moderate magnitude; hedges ES=0.53) on academic achievement in one evaluation. The IYTCM program was the only one with a positive effect on prosocial behaviour. Several programs also demonstrated positive effects on teacher outcomes such as use of praise and positive strategies (GBG and IYTCM). Importantly, these were demonstrated in RCTs of teachers managing students in the early years of school.

The evidence base for teacher training in classroom management strategies was rated Well Supported because the meta-analysis identified positive effects on relevant outcomes, had low risk of bias, and was based on studies conducted in countries considered similar to Australia. The likelihood that effects

observed in this analysis apply to the early years of schools was rated Very Plausible because contributing studies included children in the Pre-K to Grade 3 range only.

Teacher Coaching strategies

Teacher coaching strategies were identified as Well Supported strategies in two publications (Dunst, Bruder, and Hamby 2015; Kraft, Blazar, and Hogan 2018). Each is described next.

The first review included a meta-analysis of 60 studies, 56 of which were RCTs (Kraft, Blazar, and Hogan 2018). It investigated the effects of in-service teacher coaching compared with usual practice on observational measures of teachers' instructional practice and standardised measures of student achievement (in several subjects including reading, mathematics and science), across a range of grade levels from pre-K to grade 12. In this review, coaching was characterised as a process whereby instructional experts discuss classroom strategies with teachers in a way that is individualised, intensive, sustained, context-specific, and focussed. More specifically, coaching strategies in this review involved one-on-one feedback, conducted at least fortnightly, over a sustained period of time, that was specific to the teachers' own classroom and that encouraged teachers to practice specific skills. Overall, positive effects of small magnitude were observed on measures of student achievement (ES=0.18), with moderate effects on teacher instruction (ES=0.49). Coaching was more effective when paired with group training opportunities. Indeed, effects on instructional practice were 0.31 SD larger, and effects on student achievement 0.21 SD larger. Similarly, coaching combined with provision of instructional resources and materials (i.e. curriculum) was more effective than coaching without such materials. Importantly, the effects of coaching (compared with business-as-usual controls) resulted in statistically significant positive effects for children in primary school and the early years. Indeed, the effects on student achievement (across all subjects) were small but statistically significant in studies of pre-kindergarten children (ES=0.11) and primary school children (ES=0.22). Larger statistically significant effects were observed on measures of teacher instruction (including observation of pedagogical practice, teacher-student interactions, student engagement, and classroom climate) in both pre-kindergarten (ES=0.48) and primary school age children (ES=0.56).

The second relevant review was a meta-synthesis comprising fifteen research reviews including 550 studies of in-service professional development strategies to improve pre-K to 12 educator and student outcomes (Dunst, Bruder, and Hamby 2015). This synthesis specifically examined the core features of effective programs, and identified coaching as one of them. The meta-synthesis included two reviews of strategies focussing on primary school teachers, and three on early childhood educators (i.e. pre-kindergarten). While the proportion of studies within each of the reviews focussing on primary or early primary students was not clear, the authors found that the pattern of results across reviews were 'remarkably similar regardless of type of research synthesis, types of studies included in the syntheses, and types of content knowledge or practice' p.1737). The meta-synthesis found in-service professional development was effective when it included core features (described next), was of sufficient duration and intensity, and included follow-up supports to reinforce the use of content knowledge or practice. It is worth noting that these characteristics are similar to those identified in the previously described meta-analysis of coaching. Core features included: trainer introduction (of content knowledge, subject area, or practice) and illustration roles (e.g. use of modelling, simulations, etc.), active learning opportunities, and support in the form of coaching or mentoring and performance feedback during both in-service PD and follow-up.

Overall, the evidence supporting coaching strategies was rated Well Supported because at least one meta-analysis identified a positive effect on relevant outcomes, had low risk of bias, included a

sufficient proportion of primary age students, and was based on studies conducted in countries considered similar to Australia.

Strategies rated Plausible

Content-specific in-service professional development

Two publications specifically investigated the effectiveness of content-specific in-service[22] professional development for mathematics and/or science teachers.

In a report prepared for The Council of Chief State School Officers (Blank and De Las Alas 2009), a meta-analysis of professional development programs for grades K to 12 science and mathematics teachers observed an overall positive (albeit small) effect showing improved performance in mathematics (ES=0.13 to 0.21), but not science outcomes. Importantly, statistically significant effects of small-to-moderate magnitude in mathematics were also observed for studies of primary school children (ES=0.27 to 0.32). Effective programs were characterised by multiple activities to provide follow-up reinforcement of learning including coaching, mentoring, internship, professional networks, and study groups, in addition to coursework or in-service education. Active methods of teacher learning were also consistently featured in the programs. Examples include leading instruction, discussion with colleagues, observing other teachers, developing assessments, and professional networks. Common goals of the programs included improving teacher knowledge of how students learn in specific subjects, effective strategies for teaching the subject, and connection between subject content and pedagogy. Of the 16 programs, 14 had a duration of at least 6 months, and the mean contact time was 91 hours. Effect sizes were calculated for programs offering internships, using collaborative networking, and mentoring and compared with studies where PD did include these characteristics. Though a small effect on mathematics was observed when PD included internships (ES=0.20, based on nine effects), the result failed to reach significance.

A systematic review (Slavin et al. 2012) of 17 RCT and matched-control studies investigated a variety of programs and strategies used in elementary science. Enquiry-based programs emphasising professional development without the use of kits had a positive impact on student achievement in science (ES=0.30) compared with usual practice and alternative programs. The PD provided in enquiry-based programs emphasised conceptual challenges, cooperative learning, science-reading integration, teaching scientific vocabulary, and use of inquiry learning cycles.

Evidence from experimental and quasi-experimental studies

There were too few meta-analyses/systematic reviews that focused on school leadership development so an additional search was conducted to identify individual RCTs of school-based leadership strategies implemented during the primary school period (see Table 27)

[22] We define in-service professional development as training delivered to qualified and practicing teachers, rather than student-teachers (i.e. pre-service). Such training may occur at the teachers' usual school or off-site.

Table 27: Summary of CRCTs (Professional development strategies)

RCTs									
Study	**Design**	**N**	**Grades**	**Strategy**	**Outcome Area**	**Results**		**Quality**	**Evidence Rank**
(Jacob et al. 2015)	CRCT (126 schools)	Principals: 126 Teachers: 1546 Students: NR	3-5	McREL Balanced Leadership Program	- Principal efficacy School climate Collective differentiated instruction Leadership Collaboration Grade 3 Reading Grade 3 Mathematics	Principal 0.55* 0.34* 0.53* 0.33 0.40 -0.02 0.04	Teacher NA 0.02 0.07 0.03 0.06	High	Not Supported

Only one RCT investigating a leadership development program was identified. Although principal self-ratings indicate promising outcomes in terms of improved confidence, school climate, and whole school approach to differentiated instruction, teacher ratings across all measured domains and student academic outcomes suggest the program is Not Supported by the evidence. As teachers did not receive the PD, but principals did, the positive results suggested by principal self-ratings may be indicative of a social desirability bias rather than positive impact *p<0.05

Domain 11 Conclusions

The domain Teacher and Leadership Development was rated Well Supported.

Which teacher and leadership professional development strategies have demonstrated positive effects in the early years of school?

Teacher professional development strategies demonstrating positive effects on student outcomes in the early years of school include:

- Teacher training in classroom management
- In-service teacher coaching

Domain 12: Partnerships with families

Family partnership strategies aim to increase parent involvement in children's learning and development at school. Although some definitions of family partnership may include provision of intensive parent workshops or support programs implemented by specialist staff, the definition adopted here is limited to the types of strategies that schools typically have the resources to implement.

Research Questions:

12.1 Is there evidence that partnerships with families are related to student academic achievement and social, emotional, or behavioural development?

12.2 If so, which school-family partnership strategies have demonstrated positive effects for students in the early years of school?

Strength of the domain

Domain 12 was rated Supported. Correlational evidence from three meta-analyses shows parent involvement is related to student outcomes, and more critically, four meta-analyses of intervention-based research identified family-partnership strategies rated Well Supported or Supported. Moreover, the evidence shows two strategies are applicable to the early years of school. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were three correlational meta-analyses/systematic reviews identified for research question 12.1 (see Table 28 for a summary and Appendix J for details). The correlational evidence base was rated Supported for all three publications. Each provides evidence supporting an association of school-family partnerships with student academic outcomes.

There were six intervention-based meta-analyses/systematic reviews identified for research question 12.2 (see Table 30; details are presented in Appendix J). The intervention-based evidence was rated Supported for four of the publications. The types of school-family partnership strategies identified in these publications and the extent to which the evidence is relevant to the early years of school are summarised in Table 31.

Table 28: Summary of correlational meta-analyses (Family partnerships)

Study	Design	K Studies	Grades	Association	Outcome Area	Overall Association	EYS Effect (grade/age)	Quality	Evidence Rank
(Jeynes 2012)	MA	51	PK-12	Parent involvement: Shared reading Partnership Homework Communication	Academic (composite)[a]	d=0.30*, k=51 0.51* share reading 0.35* partnership 0.27* homework 0.28* communicate	Composite only: Primary d=0.29* EYS: NR	High	Supported
(Kim and Hill 2015)	MA	52	K-12	Parent involvement	Achievement (composite)[a]	r=.14* (fathers) r=.15* (mothers)	Primary r=.10*, k=12 (fathers) r=.07*, k=23 (mothers) EYS: NR	High	Supported
(Ma et al. 2016)	MA	46	K-6	Parent involvement (dimensions and types)[b]	Achievement (composite)[c]	r=0.51*	NR (though moderator analysis suggests weaker relationship in EYS)	Moderate	Supported

[a] Studies included in these meta-analyses included a range of achievement measures covering reading, language and mathematics subjects; [b] dimensions= home discussion, home supervision, home-school connection, and school participation; types= behavioural, personal, intellectual (each described more in text below); [c] the composite measure of achievement included effects from studies of achievement in language, mathematics and science subjects; NOS=not otherwise specified; *p<0.05

Correlational Evidence

Three studies were identified that Supported associations between family-school partnerships and student achievement. All investigated the association between family involvement and academic achievement (Ma et al. 2016; Jeynes 2012; Kim and Hill 2015).

Table 29 summarises associations rated Supported, along with an estimation of the likelihood that effects observed in the meta-analyses will generalise to the early years of school. The evidence base for associations rated Plausible is summarised below.

Table 29: School-family partnership association with student outcomes and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported Relationships		
None	Not applicable	Not applicable
Supported Relationships		
• Parental involvement and student achievement (Jeynes 2012; Ma et al. 2016; Kim and Hill 2015)	Small to moderate	Plausible

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific effects are presented. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small-to-Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate-to-Large=0.60 to 0.79; Large>0.80

Associations rated Plausible

Parent involvement and student academic achievement

Three meta-analyses provide evidence of an association between parent involvement and student achievement. The most recent meta-analysis conceptualised parent involvement according to two separate frameworks (Ma et al. 2016). One framework organised parent involvement in terms of four *dimensions*: home discussion (e.g. encouraging parents and students to prepare for school matters like courses, activities or events), home supervision (e.g. monitoring homework, limiting game time), home-school connection (e.g. use of communication channels between home and school), and school participation (e.g. volunteer roles at school such as serving as a reading parent or on school council or parents and friends groups). The other framework categorised family involvement according to three *types*: behavioural (e.g. visiting school and participating in educational affairs), personal (e.g. concern about affective experiences of children at school) and intellectual (e.g. reading books, solving mathematics and science problems, discussing current social and cultural events). This meta-analysis also conceptualised family-school partnerships as distinct from parental involvement. It defined family-school partnerships as recognising and emphasising 'the critical importance of open communication, healthy relationships, mutual respects (for differences), and genuine willingness to share power between families and schools.

The meta-analysis included 46 studies of grades K to 6 students. A significant positive association of moderate magnitude was reported for parental involvement and student achievement (r=0.51). Achievement included performance in language, mathematics, and science. The relationship between parental involvement and the composite measure of student achievement was statistically stronger among studies that emphasised the school participation dimension of parent involvement (compared with those that did not). However, an effect size specific to school participation studies was not reported.

The second meta-analysis supported an association between family involvement and academic achievement, which comprised 51 studies of Pre-K to 12 students (Jeynes 2012). It included a composite measure of achievement drawing on studies of various subject areas (e.g. language, reading, and mathematics) and found a small-to-moderate but statistically significant positive association (*r*=0.30). Importantly, the strength of association was analysed for several different program types including those that primarily involved (a) shared reading (i.e. programs that encourage parents and children to read together), (b) an emphasis on partnership (e.g. efforts to assist parents and teachers to collaborate as equal partners for the purposes of improving children's academic or behavioural outcomes), (c) checking homework (school-based initiatives to encourage mothers and fathers to check children complete their homework), or (d) communication between parents and teachers (e.g. efforts by schools to increase communication with parents)[23]. Statistically significant positive associations of small-to-moderate magnitude occurred for each: shared reading (*r*=0.51), emphasised partnership (*r*=0.35), checking homework (*r*=0.27), and parent-teacher communication (*r*=0.28). The meta-analysis also reported a small but statistically significant association of parental involvement with student academic achievement for students in elementary school (*r*=0.29).

A third meta-analysis investigated the association between maternal and paternal involvement with educational achievement outcomes (Kim and Hill 2015). Types of parental involvement included school-involvement, home involvement (e.g. homework assistance and intellectual enrichment), and academic socialisation. The analysis included 52 studies spanning grades K to 12. Overall, parent involvement was positively correlated with student achievement in both studies of mothers *(r*=0.15) and fathers (*r*=0.14). Small but statistically significant associations were also reported for studies of primary age students (*r*=0.07 for mothers and 0.10 for fathers). Overall, the strength of association for different types of involvement was strongest for academic socialisation. However, the strength of relationship for different forms of involvement was not reported separately for primary school students. It is also unclear to what extent the conceptualisation of family involvement in this meta-analysis is a reflection of school-family partnerships or parental interest regardless of partnership efforts. It is unclear whether included studies measured family partnerships with schools or family interest *per se*.

For all three meta-analyses the evidence was rated Supported rather than Well Supported because risk of bias was moderate and/or the countries from which the studies were drawn was not reported. Findings were considered plausible to apply to the early years of school because results specific to the primary school years were presented.

[23] The meta-analysis also investigated the relationship between parent involvement and academic achievement specifically among (a) Head Start learning centres and (b) programs where parents with a non-English speaking background were 'involved' in their child's education via school-based efforts to teach parents English.

Table 30: Summary of experimental meta-analyses and systematic reviews (Family partnerships)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Erion 2006)	MA	37 (20 group-based)	K-6	Home-based parent tutoring	Academic skills (Reading, spelling, mathematics and written expression)	0.55*,#ES=32	0.57* ,#ES=21 (grades K-3)	High	Supported
(Patall, Cooper, and Robinson 2008)	MA	14	K-12	Parent training in homework assistance	Achievement Homework completion Problems with homework Achievement only Other outcomes	RCTs 0.09, ns, k=6 0.28 [a], k=4 -1.20*, k=3 QESs 0.22*, k=3 NR	RCTs 0.23, ns, k=3 (primary achievement) EYS: No pooled ES, 3 relevant studies (mixed results).	Moderate	Supported
(See and Gorard 2015)	SR	1008 (77 rel.)	PK-12+	Parent involvement [b]	Various school outcomes [c]	EYS: Identified 3 relevant studies (2 RCT & 1 QES) suggesting positive effects but serious design flaws.		Low	Unknown
(Semke and Sheridan 2012)	SR	18	K-12	Family involvement [d]	academic, social-emotional, and behavioural outcomes	EYS: Identified only 2 relevant QESs studies		Moderate	Unknown
(Sénéchal and Young 2008)	MA	16	K-3	Parent-child reading	Literacy [e]	0.65*, k=16	0.51*, k=5 (kinder) 0.74*, k=11 (G1-3)	High	Well Supported
(Ttofi and Farrington 2011)	MA	44 (17 rel)	K-12	Anti-bullying programs (k = 17 involving parents)	Victimisation Bullying	OR: 1.41* OR: 1.57*	NR	Moderate	Supported

[a] p<0.10; [b] Parents were given and trained to use games to play with their children to develop children's pre-reading and numerical skills/ giving parents toys and books designed to encourage parent–toddler verbal interaction/ CPC which is a multi-dimensional programme which includes special classroom instruction, health advice for parents and other enrichment activities, but one of its key features is the involvement of parents in the classroom; [c] attendance, school readiness, retention, achievement; [d] parent-enriched reading groups/a parent involvement training program (no details re the training program provided); [e] word reading, early literacy skills, reading comprehension OR= Odds ratio; *p<0.05

A variety of school-family partnership strategies were identified. Well Supported strategies included involvement of parents in violence prevention and reduction programs (Ttofi and Farrington 2011). Supported strategies included home-based parent tutoring (Erion 2006), parent-child reading (Sénéchal and Young 2008), and parent training in homework assistance (Patall, Cooper, and Robinson 2008)

Table 31 summarises strategies rated Well Supported and Supported, along with an estimation of the likelihood that effects observed in the meta-analyses and systematic reviews will generalise to the early years of school. The evidence base for strategies rated Very Plausible and Plausible is summarised below.

Table 31: Effective school-family partnership strategies and EYS applicability

Overall Evidence Rating	Observed Magnitude of Effect	Likelihood that Effects Generalise to Early Years of School
Well Supported Strategies		
• Parent-child reading (Sénéchal and Young 2008)	Moderate	Very plausible
Supported Strategies		
• Inclusion of parents in violence prevention & reduction programs (Ttofi and Farrington 2011)	Small	Possible
• Home-based parent tutoring (Erion 2006)	Moderate	Very Plausible
• Parent training in homework assistance (Patall, Cooper, and Robinson 2008)	Small to large	Plausible (but note reservations)

Very Plausible: proportion of relevant studies in MA or SR focusing on student in early years of school is high (>75%) or results are presented separately for the early years of primary school. Plausible: proportion of relevant studies in MA or SR focusing on primary students is high (>75%) or primary-specific results are presented. Possible: proportion of relevant studies in MA or SR focusing on primary students is low or unclear. ES: Very Small <0.20; Small=0.20 to 0.29; Small to Moderate=0.30 to 0.49; Moderate=0.50 to 0.59; Moderate to Large=0.60 to 0.79; Large>0.80

Strategies rated Very Plausible

Home based parent tutoring

One meta-analysis investigated home-based parent tutoring. Home-based parent tutoring was defined as occurring when parents provided academic instruction to their own children at home. This meta-analysis included 20 experimental studies of students in primary school (Erion 2006). Overall, there were significant positive effects of moderate magnitude on academic achievement, measured across a range of subjects combined (ES=0.55). Importantly, a statistically significant positive effect of home-based parent tutoring of comparable magnitude was also observed for children in grades K to 3 (ES=0.57). The evidence for family-school partnerships emphasising home-based parent tutoring was rated Supported because study quality was moderate. Findings were rated Very Plausible to apply to the early years of school because the meta-analysis reported positive results specific to children in grades K to 3.

Parent-Child reading

One meta-analysis investigated the effect of school-family partnership strategies targeting parent child reading. This meta-analysis included 16 studies of grade K to 3 children including 12 experiments and four quasi-experimental designs (Sénéchal and Young 2008). Overall, parent-child reading had a positive effect of moderate-large magnitude on literacy (ES=0.65). Parent-child reading (compared with no parent reading) resulted in positive moderate-large effects for children in kinder (ES=0.51) and grades 1-3 (ES=0.74). The evidence was rated Well Supported because risk of bias was low, all studies

included primary school students, and the countries from which studies were drawn were considered similar to the Australian education systems. Findings were rated Very Plausible to apply to the early years of school because results were specific to children in grades K to 3.

Strategies rated Plausible

Parent training in homework assistance

One meta-analysis investigated parent training in homework assistance (i.e. where parents receive some type of training geared toward encouraging or improving skills for parent involvement with homework). This meta-analysis included 14 studies of students in grades K to 12, including nine experiments and five quasi-experimental designs (Patall, Cooper, and Robinson 2008). A large statistically significant effect indicated that parent training programs (compared with no parent training) led to fewer problems with homework (ES=-1.20) and a marginally significant trend indicated that training led to increases in homework completion rates (ES=0.28). Although a positive effect on achievement (mathematics and reading combined) was not significant when studies from grades K to 12 were pooled, there was a small significant effect observed for studies of children in grades 2-5 (ES=0.22).

No pooled effects were computed for studies of children in the early years of school. However, results from two relevant experimental studies were reported. Both involved the use of newsletters to encourage parental involvement in homework and included measures of achievement. Effect sizes were generally small (i.e. <0.20) and negative, but statistical significance and 95% confidence intervals were not reported. Thus, it is unclear what effect if any newsletters to encourage homework completion have on academic achievement for children in the early years of school.

Overall, the evidence for use of parent training in homework was rated Supported because the meta-analysis was of moderate quality and reported several positive effects. Findings were rated Plausible to apply to the early years of school because a statistically significant positive effect was reported at the primary school level, whereas results specific to the early years were both equivocal and limited to a specific type of parent training strategy (use of newsletters).

Domain 12 Conclusions

The domain 'family-school partnerships' was rated Supported.

Correlational evidence shows as there is an association between parental involvement and student achievement.

Which strategies targeting partnerships with families have demonstrated positive effects in the early years of school?

Specific family-school partnership strategies demonstrating positive effects for children in the early years of school include:

- Home based parent tutoring
- Parent-child reading activities

Domain 13: Community-school partnerships

Community-school partnerships involve schools collaborating with local external organisations to support student health, well-being, engagement, inclusion, or learning. Community-school partnerships may be formed with external agencies such as other schools, universities, health services and community organisations. Partnerships could range from school-based information and referral to co-location and integration of services.

Research Questions:

13.1 Is there evidence that community-school partnerships are related to student outcomes?

13.2 Which sorts of community-school partnerships have demonstrated positive effects on student outcomes in the early years of school?

Strength of the domain

Domain 13 was rated Preliminary. Only one meta-analysis, one narrative systematic review, and four quasi-experimental studies were identified. Critically, only two quasi-experimental studies presented results specific to children in grades K-3. A detailed account of the evidence is provided below.

Evidence from meta-analyses and systematic reviews

There were two meta-analyses/systematic reviews identified (see Table 32 for a summary and Appendix J for details). The evidence for strategies identified in these publications was rated Unknown as the proportion of primary age students included was either very low or not reported.

Evidence from experimental and quasi-experimental studies

Four quasi-experimental studies were identified in the secondary search (see Table 33 for a summary and Appendix J for details). Each of the four studies investigated the effect of a community-school partnership with at least one external organisation and reported positive effects on student outcomes. There were however no Well Supported or Supported experimental studies identified. As such, the evidence was rated Preliminary.

Table 32: Summary of meta-analyses and systematic reviews (Community-school partnerships)

Study	Design	K Studies	Grades	Strategy	Outcome Area	Overall Effect	EYS Effect (grade/age)	Quality	Evidence Rank
(Celio, Durlak, and Dymnicki 2011)	MA	62	K-12+	Service-learning[a]	Academic Attitude to self Attitude to school Civic engagement Social skills	0.43*, k=17 0.28*, k=35 0.28*, k=12 0.27*, k=28 0.30*, k=28	NR (5% primary)	Moderate	Unknown
(Davies et al. 2013)	SR	58 (9 rel.)	K-12	Conditions for creativity (use of environments beyond the school) [b]	Creative skills	9 studies [c] suggest involvement with external agencies supports a creative environment	NR	Moderate	Unknown

[a] Service learning was defined as community service integrated with academic curriculum [b] Examples of involvement with external agencies included visiting museums and galleries, youth clubs, and engaging with local businesses, and the wider sports and arts communities. Of the nine studies cited, only four were described and these appear limited to case studies of secondary education initiatives

Table 33: Summary of QES investigations (Community-school partnerships)

Experimental and Quasi-Experimental Community-School Partnership Studies								
Study	**Design**	**N**	**Grades**	**Strategy**	**Outcome Area**	**Results**	**Quality**	**Evidence Rank**
(Henry, Bryan, and Zalaquett 2017)	QES	Teachers: NR Students: 1290	3-5	Just Love [a]	Reading	Mixed. Small Positive effects at school level in 2 of 3 implementation years. Effects at classroom and student level, ns.	Moderate	Preliminary
(Hoglund et al. 2012)	QES (17 schools)	Teachers: n = ~30 IG; ~16 CG) Students : 432	1-3	WITS [b]	Physical victimisation Relational victimisation Social competence Physical aggression Help-seeking Internalising	0.17* 0.20* 0.20* 0.09* 0.04 0.10	High	Preliminary
(Lee and Stewart 2013)	QES (20 schools)	Teachers: NR Students : 2758)	3,5,7	Health Promoting School partnership [c] EYS: effects not reported separately but mean age 10 years	Family connection * School connection Community connection* Peer support* Resilience* Autonomy	Overall: +ve, sig. on 4 outcomes (effect sizes NR)	Moderate	Preliminary
(Weaver et al. 2018)	QES (3 schools)	Teachers : 15 Students : 229	1-3	PACES [d]	% children receiving 30 minutes MVPA over day % time in Moderate Vigorous Physical Activity daily	Girls: ns Boys: ns Boys: 2.13* Girls: 0.70*	Moderate	Preliminary

[a] Just Love is a counsellor-led, faith-based, school-family partnership including student mentoring, volunteer teaching aide, and school-wide incentive and enrichment program, [b] WITS is a Community based, whole school peer victimisation prevention program that teaches children the following strategies for coping with bullying: Walk away, Ignore, Talk, Seek help., [c] Schools partnered with local communities and various organizations (e.g. local city council, local Departments of Health and Education, and NGOs) [d] Partnerships for Active Children in Elementary School - A school-university partnership utilising communities of practice, service-learning, and community-based participatory research approaches to partnership

Note: In domains where Supported and Well Supported strategies have been identified, only these have received Plausibility ratings. However, given the relative dearth of good quality trials in this domain, an applicability rating and brief description is provided for strategies rated Preliminary.

Preliminary Strategies rated Very Plausible

The results for two of the community partnership strategies (WITS and PACES) were rated Very Plausible to apply to the early years of school. Each is described next.

WITS

WITS is a community-based, whole-school peer victimisation prevention program (Hoglund et al. 2012). The program name is derived from several strategies taught to reduce risk of peer victimisation. These are "Walk away (and seek help), Ignore it (and seek help), Talk it out (and seek help), and Seek help". While the program is delivered by teachers, it was developed as a collaborative effort between school staff members, a community-based police group, the Rock Solid Foundation, and developmental psychologists. The program has been tested in one quasi-experimental study including seventeen schools (Hoglund et al. 2012). The trial included 432 students (290 intervention vs 142 control) in grade 1. The program demonstrated small but significant positive changes on measures of physical victimisation (ES=0.17), relational victimisation (ES=0.20), social competence (ES=0.20) and physical aggression (ES=0.09). There were no significant effects for help seeking behaviours or internalising problems. The evidence for this strategy was rated Preliminary because positive effects were observed in a reasonably well-conducted quasi-experimental study. The findings are considered Very Plausible to apply to the early year of school because positive effects were demonstrated for students in grade 1.

Partnerships for Active Children in Elementary School (PACES)

PACES is a school-university partnership utilising collaborative approaches that provide schools with external support for implementing physical activity promotion strategies; communities of practice, service-learning, and community-based participatory research approaches to partnership (CBPR)(Weaver et al. 2018). Consistent with CBPR principles of engaging community partners the lead author met with PE teachers to orient them to the intervention and set goals for increasing children's moderate-vigorous physical activity (MVPA) during lessons. Researchers also met with each classroom teacher to introduce a free online virtual communities of practice website designed to help classroom teachers integrate movement into classroom time. Service-learning involved sending preservice teachers into classrooms to deliver movement integration strategies between academic lessons.

The effect of PACES on moderate-vigorous physical activity (MVPA) was tested in a quasi-experimental study involving three schools (Weaver et al. 2018). This trial included 229 students in grades 1 to 3 and 15 teachers. Results indicated that although there was no significant difference in the percentage of intervention and comparison children accumulating 30 minutes of MVPA over the course of an entire school day, there was a significant positive effect of moderate-large magnitude in the percentage of time spent in MVPA (this was reported separately for boys and girls, not combined). The effect was larger for boys (ES=2.13) than girls (ES=0.70). There was also a significant and positive effect of the program on the percentage of time spent in MVPA during physical education classes, though this was significant for boys only (ES=0.73). No significant treatment effects were observed on the percentage of time spent in sedentary activity.

The evidence for this strategy was rated Preliminary because positive effects were observed in a moderate quality quasi-experimental study. Findings were rated Very Plausible to apply to the early years of school because for the program was tested with (and results reported for) children in grades 1 to 3.

Preliminary Strategies rated Plausible

The results for two other community partnership strategies (Health Promoting School Partnerships and Just Love) were rated plausible to apply to the early years of school. Each is described next.

Health Promoting School Partnership

The Health Promoting School (HPS) Partnership is described as a holistic, multi-strategy model including: constant communication and shared visions; staff empowerment; providing a structure that supports a culture of HPS; and support for school partnerships with families and communities in an effort to build student resilience (Lee and Stewart 2013). Schools were connected with local communities and various organisations such as local city council, local Departments of Health and Education, and NGOs, which provided the school with a range of support services and resources. Specific types of support and services were not described.

A quasi-experimental study of the model involving twenty Queensland primary schools investigated potential HPS effects on student resilience following 18-months of implementation (Lee and Stewart 2013). This trial included 2,758 students in grades 3, 5 and 7. Statistically significant positive program effects were observed for student-family connection, student-community connection, student-peer connection, and student resilience (effect sizes not reported). Although results were not presented separately by grade level, results suggested a protective effect of HPS on resilience was stronger for younger students. The evidence for the HPS partnership model was rated Preliminary because positive effects were observed in a moderate quality quasi-experimental study. The extent to which findings can be applied to the early years of school is somewhat unclear. While it seems likely most students would have been primary school age, the proportion is not reported and nor were results presented separately for children in primary school or the early years of school specifically. As the mean age of participants was 10 years, however, it seems plausible that results might generalise to the early years of school.

Just Love

Just Love is a counsellor-led, faith-based school-family-community partnership comprising three component programs (Henry, Bryan, and Zalaquett 2017). Component programs include Just Mentor (where students with significant academic or behavioural problems are mentored by a volunteer for a full year), Just Connect (where a small group of volunteers serve as teaching aides and provide classroom-based small group and individual tutoring), and Just Rewards (a school-wide incentive and enrichment program involving rewards to encourage regular and punctual attendance, improved behaviour and academic achievement). Examples of enrichment opportunities include lessons in ballet or karate, parent workshops, and community fairs. The program was developed by a partnership leadership team comprising a school principal, school counsellor, other student service personnel, and a parent. It is important to note that faith-based volunteers delivering the program were trained in school policies and procedures, and were instructed to adhere to guidelines explicating the importance of not advocating particular religious or political viewpoints.

The potential effect of Just Love on reading achievement was tested in a quasi-experimental study over three school years 2010–2013 (Henry, Bryan, and Zalaquett 2017). The study included 1290 students in grades 3 to 5. Statistically significant differences were observed indicating that students at the intervention school outperformed students at a comparison school on a standardised measure of reading achievement in two out of three implementation years (not the first year of implementation). However, comparisons between intervention and comparison classrooms, and mentored and non-mentored students were not significant. The evidence for this strategy was rated Preliminary because some positive effects were observed in a moderate quality quasi-experimental study. Findings are

considered plausible to apply to the early years of school because positive effects were observed for primary school students.

Domain 13 Conclusions

The evidence base for the domain community-school partnerships was rated Preliminary.

One meta-analysis of intervention research reported positive effects of service learning on academic achievement, civic engagement, social skills, and positive attitudes toward self and school.

Which community-school partnership strategies have demonstrated positive effects in the early years of school?

None of the identified strategies have been rigorously tested in RCTs with children in the early years of school.

CONCLUSION

Summary

Although there are existing school quality frameworks in use, the evidence base underpinning these frameworks is often not well-documented or rigorously and consistently evaluated.

The aim of this restricted review was to: (a) evaluate the evidence base for 13 school quality domains commonly identified in existing school quality frameworks, (b) identify the specific strategies within each domain that are both supported by a strong evidence base and have demonstrated effectiveness in improving child academic and psychosocial outcomes in the early years of school, and (c) develop an evidence-based list of quality indicators for use within school improvement assessments. The review included 83 relevant publications, of which 66 were meta-analyses or reviews, providing the highest-levels of evidence. An evaluation of the evidence-base was conducted for each of the 13 domains, identifying 5 Well Supported and 4 Supported quality domains. Within these domains, the review identified 21 general strategies that have demonstrated effectiveness for children in the early years of school (see Table 34). These findings informed the development of an evidence-based framework of indicators for establishing school quality.

Table 34: Evidence-based strategies to improve child academic and psychosocial development

Domains	Overall Domain Evidence Rating	Strategies Applicable to the Early Years of School
Part I: Effective Classroom Teaching Practices		
1. Application of pedagogical content knowledge	Well Supported	• Explicit teaching of phonemic awareness • Explicit teaching of phonics, fluency, and reading comprehension strategies • Balanced reading and writing instructional time • Explicit morphological instruction • Directly teaching spelling skills • Use of musical context to teach reading and literacy skills • Provision of explicit handwriting instruction • Use of manipulatives in mathematics
2. Effective differentiated teaching strategies	Well Supported	• Literacy instruction in small groups • Computerised instructional differentiation
3. Peer teaching and co-operative learning approaches	Supported	• Peer teaching (evidence for specific strategies unclear)
4. Use of physical activity	Supported	Review level evidence for strategies not reported separately for grades K-3
5. Technology-assisted teaching and learning	Well Supported	• ABRACADABRA (supplementary, interactive, web-based reading skills program) • Exploratory digitised environments in mathematics instruction • Various technological tools such as computers, interactive whiteboards, and multi-media (used interactively)
6. Physical environment design to optimise learning	Preliminary	Evidence limited for EYS
7. Class size and Teacher-Student ratios	Supported	• 22 students or fewer per class
Part II: The School Culture		
8. Student empowerment and leadership	Unknown	No evidence-based student empowerment and leadership strategies identified

Domains	Overall Domain Evidence Rating	Strategies Applicable to the Early Years of School
9. Social-emotional and behavioural interventions to promote a positive school climate	Well Supported	• Teacher training in classroom management (e.g. Incredible Years Teacher Classroom Management Program; Good Behaviour Game) • Mindfulness programs
10. Teacher-student relationships	Preliminary	Evidence limited for EYS
Part III: Providers and Partnerships		
11. Staff and leadership development	Well Supported	• In-service teacher coaching (that is individualised, intensive, sustained, context-specific, and focussed) • Classroom management training
12. Partnerships with families	Supported	• Home-based parent tutoring • Parent-child reading activities
13. Community-school partnerships	Preliminary	Evidence limited for EYS

EYS quality indicators

Quality indicators were informed by:

1. Identification of Supported or Well Supported school quality domains originally identified in existing frameworks – that is, domains for which at least one effective strategy is supported by at least one good quality meta-analysis or systematic review (and its use is not contra-indicated by a similar quality meta-analysis or systematic review).

2. Evaluation of the extent to which the evidence underpinning Supported and Well Supported domains is applicable to the EYS.

In total, 37 quality indicators were developed. For pragmatic purposes, these indicators are tied to school processes and teaching staff competencies that map to Well Supported and Supported quality domains. Indicators relating to Well Supported domains are presented in Table 35 and those for Supported domains appear in Table 36 (see Appendix L for definitions). Together with the indicators, the evidence-based strategies that have demonstrated effectiveness for the EYS provide a useful resource for guiding school selection of quality improvement initiatives.

Table 35: School Quality Indicators – Well Supported Domains

	Process	Provider
Content Knowledge	**QL 1** % of K-3 classroom teachers who utilise the school curriculum to plan pedagogical content delivery **QL 2** % of K-3 classrooms that balance the amount of time spent in reading and writing activities **QL 3** % of K-3 classrooms implementing daily literacy instruction that explicitly builds skills in phonics, phonemic awareness, spelling, morphology, reading fluency and comprehension strategies, and handwriting **QL 4** % of K-3 classrooms that incorporate regular use of manipulatives in numeracy instruction	**QL 5** % of K-3 classroom teachers who have formal training in evidence-based teaching methods **QL 6** % of K-3 classroom teachers who have formal training in evidence-based teaching methods who regularly coach other staff delivering K-3 literacy and numeracy
Differentiated Teaching	**QL 7** % of K-3 students whose academic development in literacy and numeracy is systematically assessed and documented **QL 8** % of K-3 students whose literacy and numeracy instruction is tailored according to the results of systematic assessment of their academic development **QL 9** % of K-3 students who regularly receive instruction in small groups **QL 10** % of K-3 students for whom assessment data indicates the need for individualised instruction in literacy or numeracy who receive an evidence based Tier 3 intervention	**QL11** % of K-3 classroom teachers with formal training in evidence-based differentiated teaching strategies **QL 12** % of staff with formal training or tertiary qualifications in special education for K-3 students needing additional support **QL 13** % of staff delivering additional support to K-3 students who have formal training in the provision of evidence based Tier 2 and Tier 3 learning interventions

	Process	Provider
Technology	QL 14 % of K-3 lessons utilising digital technology for instruction in interactive rather than static conditions QL 15 % of K-3 classrooms utilising interactive digital technology platforms to supplement literacy and numeracy instruction	QL 16 % of classroom teachers who have received formal training in the use of interactive digital instruction materials and incorporate these in their classes
Social Emotional Support	QL 17 An evidence-based social-emotional development program is implemented across the school and activities to maintain the skills developed in the program are delivered on a regular basis (i.e. every term) QL 18 % of families (with a child in grade K-3) indicating that their child feels safe at school on annual parent surveys QL 19 % of families (with a child in grade K-3) who agree on parent opinion surveys that teachers at the school treat students fairly and/or student behaviour is well managed	QL 20 % of K-3 classroom teachers who have completed formal training in evidence-based social-emotional development programs (such as teaching mindfulness strategies)
Staff Development & Leadership	QL 21 % of staff professional development opportunities approved by the school principal that are characterised by both (a) active teacher learning experiences and (b) use of modelling/simulations QL 22 % of approved professional development opportunities that are informed by student needs (I.e. based on data)	QL 23 % of K-3 classroom teachers with formal training in an evidence-based classroom management strategy QL 24 % of professional learning courses undertaken by teachers that are evidence-based QL 25 % of teachers currently receiving in-service teacher coaching that is considered best practice

Table 36: School Quality Indicators – Supported Domains

	Process	Provider
Peer Teaching	QL 26 % of K-3 classrooms that implement evidence-based peer tutoring activities in the weekly literacy/numeracy blocks	QL 27 % of K-3 classroom teachers with formal training in evidence-based peer teaching methods
Physical Activity	QL 28 % of K-3 classrooms where physical activity is incorporated in academic instruction on a daily basis (whether by in class activity breaks, exercise prior to lessons, or use of movement to facilitate instruction)	QL 29 % of K-3 classroom teachers who have received at least some informal training in strategies to incorporate movement in academic instruction
Class Size	QL 30 % of K-3 classes that comprise 22 students or fewer	

	Process	Provider
Partnerships with Families	**QL 31** % of K-3 teachers who are aware of the school's family partnership policy and implement it into their usual practice with families **QL 32** % of families (with a child in grades K-3) indicating that the school actively encourages and emphasises the importance of regular parent-child reading at home **QL 33** % of families (with a child in grades K-3) indicating that the school has provided information about specific strategies for parents to use when reading with their children	**QL 34** % of K-3 classroom teachers indicating that they have provided parents with strategies to use when reading with children at home **QL 35** % of K-3 classroom teachers indicating that they monitor parent home reading on a regular basis (i.e. weekly) **QL 36** % of K-3 classroom teachers indicating that they provide additional support to parents who have difficulties with home reading practice **QL 37** % of K-3 classroom teachers indicating that the materials they provide parents (to encourage and support reading at home) are evidence-based

Strengths of the approach

This restricted review focussed on studies utilising the most rigorous methods of evaluation (primarily meta-analyses and systematic reviews) to provide the strongest levels of evidence in identifying effective EYS practices. The review spanned a wide range of quality domains and child outcomes informed by existing, contemporary, and predominantly Australian school quality frameworks. As such the resulting school quality indicators should be both useful as stand-alone measures and easily relatable to any existing frameworks. In contrast to existing frameworks, this review also explicitly evaluates the evidence for each domain, using a consistent approach that allows comparison of the strength of evidence for each. Moreover, the estimation of applicability to the early years of school provides an indication of which domains and strategies are most suitable for early intervention. Finally, as the search was conducted across several of the most relevant academic databases and a dual-phase approach to searching was utilised, the search strategy ensured extensive coverage of relevant publications. The resulting school quality indicators will be useful as stand-alone measures, while also being easily relatable to any existing frameworks used in schools.

Limitations of the approach

The broad scope of included domains necessitated some concession to the breadth and depth of the review. In cases where the sheer volume of potentially relevant literature exceeded resource capacity and multiple recent meta-analyses or systematic reviews met inclusion criteria for a given domain, additional selection restrictions were applied (e.g. further publication date limits were applied for the domains of pedagogical content knowledge, and digital technology). With respect to review depth, it is important to note that in cases where multiple meta-analyses or systematic reviews met inclusion criteria and identified at least one supported strategy, searches for lower-levels of evidence were not conducted. It is possible that, in these instances, the review has missed some promising strategies. However, the purpose of the review was to identify those domains and strategies with the strongest levels of evidence rather than to provide a comprehensive list of potentially effective strategies.

Gaps in the literature and directions for future research

Low levels of evidence for several of the review domains point to major gaps in the research literature (rather than to strong evidence of ineffectiveness or detrimental effects). There is a need for future

investigations to rigorously evaluate the following domains: physical environment to optimise learning; student empowerment and leadership development; teacher-student relationships; community-school partnerships). This is important because it is possible that some of the identified strategies for these domains may in fact be effective. However, the effectiveness of such strategies cannot be established with any certainty until rigorously evaluated.

Within several domains, there are also important gaps in the literature pertaining to specific strategies. For example, even though the peer-teaching domain was rated Supported and moderately applicable to the early years of school, none of the meta-analyses investigated which specific peer-teaching strategies (e.g. ability-matched pairing; group-based vs individual reward contingencies etc.) were most effective for students in the early years of school. Similarly, there is a dearth of research evidence assessing the long-term benefits of specific strategies. Very few reviews or meta-analyses were able to include evaluation of maintenance effects.

Finally, very little is known about the extent to which the identified strategies are effective for children from disadvantaged backgrounds. The proportion of low SES samples was reported for only 14 of the strategies rated Well Supported or Supported and Very Plausible or Plausible to apply to the early years of school. In only three of these was the proportion of studies with predominantly low SES samples 50% or more (see Appendix K). Thus, there is a need for future research to evaluate whether the effectiveness of the identified strategies is moderated by SES.

Implications

Overall, the review indicates that there is a reasonably strong evidence base supporting several of the domains identified in existing school quality frameworks. Indeed, nine of thirteen identified domains were rated Supported or Well Supported. The review also shows many of the strategies underpinning these domains have demonstrated effectiveness for children in the early years of school.

The results of this review have informed the development of key school quality indicators based on a robust evaluation of the evidence that can be used to assess, monitor, and guide school quality improvement initiatives. The preliminary indicators we have selected will help identify gaps and priorities for Australian schools. We will test them in 8-10 communities over the next three years to determine which are pragmatic to collect, resonate with school communities, and provide robust measures to stimulate community and government action. We will follow a similar path for the other four fundamental strategies that Restacking the Odds is focusing on – antenatal care, sustained nurse home visiting, parenting programs, and early childhood education and care.

REFERENCES

Abrami, P.C., R.M. Bernard, E. Borokhovski, D.I. Waddington, C.A. Wade, and T. Persson. 2015. 'Strategies for teaching students to think critically: A meta-analysis', *Review of Educational Research*, 85: 275-314.

Abrami, P.C., E. Borohkovski, and L. Lysenko. 2015. 'The effects of ABRACADABRA on reading outcomes: A meta-analysis of applied field research', *Journal of Interactive Learning Research*, 26: 337-67.

Abry, T., S.E. Rimm-Kaufman, R.A. Larsen, and A.J. Brewer. 2013. 'The influence of fidelity of implementation on teacher-student interaction quality in the context of a randomized controlled trial of the Responsive Classroom approach', *Journal of School Psychology*, 51: 437-53.

ACARA. 2016. "Year 12 certification rates by geolocation and sex, by state/territory, 2018 (per cent)." In.

———. 2018. "NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2018." In. Sydney.

Amlani, A.M., and T.A. Russo. 2016. 'Negative effect of acoustic panels on listening effort in a classroom environment', *Journal of the American Academy of Audiology*, 27: 805-15.

ANZHES. n.d. *Dictionary of Educational History in Australia and New Zealand* (Austalian and New Zealand History of Education Society).

Australian Children's Education & Care Quality Authority. 2009. "National Quality Standard for Early Childhood Education and Care and School Aged Care." In. Sydney, Australia: Australian Children's Education & Care Quality Authority.

Barbero, J.A.J., J.A.R. Hernández, B.L. Esteban, and M.P. García. 2012. 'Effectiveness of antibullying school programmes: A systematic review by evidence levels', *Children and Youth Services Review*, 34: 1646-58.

Barnett, S.W. . 2011. 'Effectiveness of early educational intervention', *Science*, 33: 975-78.

Bennett, John, and Collette P Tayler. 2006. "Starting strong II: Early childhood education and care." In. Paris, France: OECD.

Blank, R.K., and N. De Las Alas. 2009. *The Effects of Teacher Professional Development on Gains in Student Achievement: How Meta Analysis Provides Scientific Evidence Useful to Education Leaders* (The Council of Chief State Officers: Washington, DC).

Bope, Rebecca, and Lennie Barblett. 2016. 'The NQS goes to school', *Every Child*, 22: 36-37.

Cappella, E., B.K. Hamre, H.Y. Kim, D.B. Henry, S.L. Frazier, M.S. Atkins, and S.K. Schoenwald. 2012. 'Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools', *Journal of Consulting and Clinical Psychology*, 80: 597-610.

Carbonneau, K.J., S.C. Marley, and J.P. Selig. 2013. 'A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives', *Journal of Educational Psychology*, 105: 380-400.

Carlisle, B.L. , and C.B. Murray. 2015. 'Academic performance, effects of socio-economic status on', *International Encyclopedia of the Social & Behavioral Sciences 2nd Edition*, 1: 43-48.

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

CEBC. 2016. "The California Evidence-based Clearinghouse for Child Welfare Scientific Rating Scale." In.: California Department of Social Services.

Celio, C.I., J. Durlak, and A. Dymnicki. 2011. 'A meta-analysis of the impact of service-learning on students', *Journal of Experiential Education*, 34: 164-81.

Cheung, A.C.K., and R.E. Slavin. 2012. 'How features of educational technology applications affect student reading outcomes: A meta-analysis', *Educational Research Review*, 7: 198-215.

Chung, K.K.H. . 2015. 'Socioeconomic status and academic achievement', *International Encyclopedia of the Social & Behavioral Sciences 2nd Edition*, 22: 924-30.

Cohen, A.K., and S.L. Synne. 2013. 'Education: A missed opportunity for public health intervention', *American Journal of Public Health* e1-e5.

Cornelius-White, J. 2007. 'Learner-centered teacher-student relationships are effective: A meta-analysis', *Review of Educational Research*, 77: 113-43.

CSHD. 2008. "CLosing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health." In.

Davies, D., D. Jindal-Snape, C. Collier, R. Digby, P. Hay, and A. Howe. 2013. 'Creative Learning Environments in Education--A Systematic Literature Review', *Thinking Skills and Creativity*, 8: 80-91.

Dennis, M.S., E. Sharp, J. Chovanes, A. Thomas, R.M. Burns, B. Custer, and J. Park. 2016. 'A Meta-Analysis of Empirical Research on Teaching Students with Mathematics Learning Difficulties', *Learning Disabilities Research & Practice (Wiley-Blackwell)*, 31: 156-68.

Department of Education and Training. 2019. "Framework for Improving Student Outcomes (FISO)." In.: State Goverment of Victoria Australia.

Deunk, M.I., A.E. Smale-Jacobse, H. de Boer, S. Doolaard, and R.J. Bosker. 2018. 'Effective differentiation practices: A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education', *Educational Research Review*, 24: 31-54.

Dietrichson, J., M. Bøg, T. Filges, and A.K. Jørgensen. 2017. 'Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis', *Review of Educational Research*, 87: 243-82.

Donnelly, J.E., J.L. Greene, C.A. Gibson, B.K. Smith, R.A. Washburn, D.K. Sullivan, K. DuBose, M.S. Mayo, K.H. Schmelzle, and J.J. et al. Ryan. 2009. 'Physical Activity Across the Curriculum (PAAC): a randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children', *Preventive medicine*, 49: 336-41.

Donnelly, J.E., C.H. Hillman, J.L. Greene, D.M. Hansen, C.A. Gibson, D.K. Sullivan, J. Poggio, M.S. Mayo, K. Lambourne, and A.N. et al. Szabo-Reed. 2017. 'Physical activity and academic achievement across the curriculum: results from a 3-year cluster-randomized trial', *Preventive medicine*, 99: 140-45.

Dunst, C.J., M.B. Bruder, and D.W. Hamby. 2015. 'Metasynthesis of in-service professional development research: Features associated with positive educator and student outcomes', *Educational Research and Reviews*, 10: 1731.

Durlak, J.A., R.P. Weissberg, A.B. Dymnicki, R.D. Taylor, and K.B. Schellinger. 2011. 'The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions', *Child Development*, 82: 405-32.

Ehri, L.C., S.R. Nunes, S.A. Stahl, and D.M. Willows. 2001. 'Systematic Phonics Instruction Helps Students Learn to Read: Evidence from the National Reading Panel's Meta-Analysis ', *Review of Educational Research*, 71: 393-447.

Elbaum, B., S. Vaughn, M. Hughes, and S.W. Moody. 1999. 'Grouping practices and reading outcomes for students with disabilities', *Exceptional Children*, 65: 399-415.

Elleman, A.M. 2017. 'Examining the impact of inference instruction on the literal and inferential comprehension of skilled and less skilled readers: A meta-analytic review', *Journal of Educational Psychology*, 109: 761.

Erion, J. 2006. 'Parent tutoring: A meta-analysis', *Education & Treatment of Children*, 29: 79-106.

Fernandez, M.A., J.S. Adelstein, S.P. Miller, M.J. Areizaga, D.C. Gold, A.L. Sanchez, S.A. Rothschild, E. Hirsch, and O.G. Gudino. 2015. 'Teacher-child interaction training: A pilot study with random assignment', *Behavior Therapy*, 46: 463-77.

Ferraro, K.F., and T.P. Shippee. 2009. 'Aging and cumulative inequality: How does inequality get under the skin?', *The Gerontologist*, 49: 333-43.

Fisher, A.V., K.E. Godwin, and H. Seltman. 2014. 'Visual environment, attention allocation, and learning in young children: When too much of a good thing may be bad', *Psychological Science*, 25: 1362-70.

French, M.T., J.F. Homer, I. Popovici, and P.K. Robins. 2015. 'What you do in high school matters: High school GPA, educational attainment, and labor market earnings as a young adult', *Eastern Economic Journal*, 41: 370-86.

Glass, G.V., and M.L. Smith. 1979. 'Meta-analysis of research on class size and achievement', *Educational evaluation and policy analysis*, 1: 2-16.

Goodwin, A.P., and S. Ahn. 2013. 'A meta-analysis of morphological interventions in english: Effects on literacy outcomes for school-age children', *Scientific Studies of Reading*, 17: 257-85.

Gordon, R.L., H.M. Fehd, and B.D. McCandliss. 2015. 'Does music training enhance literacy skills? A meta-analysis', *Frontiers in Psychology Vol 6 2015, ArtID 1777*, 6.

Graham, S., M. Hebert, and K.R. Harris. 2015. 'Formative assessment and writing: A meta-analysis', *Elementary School Journal*, 115: 523-47.

Graham, S., and T. Santangelo. 2014. 'Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review', *Reading and Writing: An Interdisciplinary Journal*, 27: 1703-43.

Graham, S., L. Xinghua, A. Aitken, C. Ng, B. Barlett, K.R. Harris, and J. Holzapfel. 2018. 'Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis', *Reading Research Quarterly*, 53: 279-304.

Gunter, T., and J. Shao. 2016. 'Synthesizing the effect of building condition quality on academic performance', *Education Finance and Policy*, 11: 97-123.

Hahn, R.A., and B.I. Truman. 2015. 'Education improves public health and promotes health', *International journal of health services*, 45: 657-78.

Hammill, D.D., and H. Swanson. 2006. 'The national reading panel's meta-analysis of phonics instruction: Another point of view', *The Elementary School Journal*, 107: 17-26.

Hard, Louise, and Lyndal O'Gorman. 2007. ''Push-Me'or 'Pull-You'? An Opportunity for Early Childhood Leadership in the Implementation of Queensland's Early Years Curriculum', *Contemporary Issues in Early Childhood*, 8: 50-60.

Harrison, Linda J, Judy Ungerer, Grant J Smith, Stephen R Zubrick, Sarah Wise, Frances Press, and Manjula Waniganayake. 2010. "Child care and early education in Australia-The longitudinal study of Australian children." In. Canberra, Australia: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

Harvey, S., K. Lambourne, J.L. Greene, C.A. Gibson, J. Lee, and J.E. Donnelly. 2018. 'The Effects of Physical Activity on Learning Behaviors in Elementary School Children: a Randomized Controlled Trial', *Contemporary School Psychology (Springer Science & Business Media B.V.)*, 22: 303-12.

Henry, L.M., J. Bryan, and C.P. Zalaquett. 2017. 'The Effects of a Counselor-Led, Faith-Based, School-Family-Community Partnership on Student Achievement in a High-Poverty Urban Elementary School', *Journal of Multicultural Counseling & Development*, 45: 162-82.

Hoffman, J.I.E. 2015. 'Meta-analysis.' in, *Biostatistics for Medical and Biomedical Practitioners* (Academic Press).

Hoglund, W.L., G. Hosan, E. Naheed , and B.J. Leadbeater. 2012. 'Using Your WITS: A 6-Year Follow-Up of a Peer Victimization Prevention Program', *School Psychology Review*, 41: 193-214.

Jacob, R., R. Goddard, M. Kim, R. Miller, and Y. Goddard. 2015. 'Exploring the Causal Impact of the McREL Balanced Leadership Program on Leadership, Principal Efficacy, Instructional Climate, Educator Turnover, and Student Achievement', *Educational Evaluation and Policy Analysis*, 37: 314-32.

James, E., A. Freund, A. Booth, M.J. Duncan, N. Johnson, and C.E. Short. 2016. 'Comparative efficacy of simultaneous versus sequential multiple health behavior change interventions among adults: A systematic review of randomised trials.', *Prev Med*, 89: 211-23.

Jay, Jenny, Marianne Knaus, and Sandra Hesterman. 2014. 'High-quality early childhood education in the early years of school', *Every Child*, 20: 22.

Jeynes, W. 2012. 'A meta-analysis of the efficacy of different types of parental involvement programs for urban students', *Urban Education*, 47: 706-42.

Kim, S., and N.E. Hill. 2015. 'Including Fathers in the Picture: A Meta-Analysis of Parental Involvement and Students' Academic Achievement', *Journal of Educational Psychology*, 107: 919-34.

Korpershoek, H., T. Harms, H. de Boer, M. van Kuijk, and S. Doolaard. 2016. 'A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes', *Review of Educational Research*, 86: 643-80.

Kraft, M.A., D. Blazar, and D. Hogan. 2018. 'The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence', *Review of Educational Research*, 88: 547-88.

Krieg, Susan, and Kay Whitehead. 2015. 'Where are the early years of school in contemporary early childhood education reforms? An historical perspective', *The Australian Educational Researcher*, 42: 319-33.

Kunsch, C.A., A.K. Jitendra, and S. Sood. 2007. 'The effects of peer-mediated instruction in mathematics for students with learning problems: A research synthesis', *Learning Disabilities Research & Practice*, 22: 1-12.

Ladd, H., and S. Loeb. 2013. 'The Challenges of Measuring School Quality: Implications for Educational Equity.' in D. Allen and R. Reich (eds.), *Education, Justice, and Democracy* (University of Chicago Press: Chicago, IL).

Laevers, Ferre. 2005. 'The curriculum as means to raise the quality of early childhood education. Implications for policy', *European Early Childhood Education Research Journal*, 13: 17-29.

Lamb, S., J. Jackson, A. Walstab, and S. Huo. 2015. "Educational opportunity in Australia 2015: Who succeeds and who misses out." In.: Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute, Melbourne: Mitchell Institute.

Lee, P.C., and D.E. Stewart. 2013. 'Does a Socio-Ecological School Model Promote Resilience in Primary Schools?', *Journal of School Health*, 83: 795-804.

Lee, Y.H., H. Waxman, J.Y. Wu, G. Michko, and G. Lin. 2013. 'Revisit the effect of teaching and learning with technology', *Educational Technology & Society*, 16: 133-46.

Leflot, G., P. van Lier, P. Onghena, and H. Colpin. 2010. 'The role of teacher behaviour management in the development of disruptive behaviours: an intervention study with the good behaviour game.', *Journal of Abnormal Child Psychology.*, 38: 869-82.

Lei, H., Y. Cui, and M.M. Chiu. 2016. 'Affective teacher-student relationships and students' externalizing behavior problems: A meta-analysis', *Frontiers in Psychology Vol 7 2016, ArtID 1311*, 7.

Leithwood, K., and J. Sun. 2012. 'The Nature and Effects of Transformational School Leadership: A Meta-Analytic Review of Unpublished Research', *Educational Administration Quarterly*, 48: 387-423.

Leung, K.C. 2015. 'Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement', *Journal of Educational Psychology*, 107: 558-79.

Levin, H., C. Belfield, P. Muennig, and C. Rouse. 2007. "The costs and benefits of an excellent education for all of America's children." In.

Li, Y., D. Coleman, M. Ransdell, L. Coleman, and C. Irwin. 2014. 'The Effect of a Sport Stacking Intervention on Handwriting with Second Grade Students', *Physical Educator*, 71.

Lim, P., S. Gemici, and T. Karmel. 2014. 'The impact of school academic quality on low socioeconomic status students', *The Australian Economic Review*, 47: 100-06.

Lochner, L. , and E. Moretti. 2004. 'The effect of education on crime: Evidence from prison inmates, arrests, and self-reports.', *American Economic Association*, 94: 155-89.

Lou, Y., P.C. Abrami, and J.C. Spence. 2000. 'Effects of within-class grouping on student achievement: An exploratory model', *The Journal of Educational Research*, 94: 101-12.

Ma, X., J.P. Shen, H.Y. Krenn, S.S. Hu, and J. Yuan. 2016. 'A Meta-Analysis of the Relationship Between Learning Outcomes and Parental Involvement During Early Childhood Education and Early Elementary Education', *Educational Psychology Review*, 28: 771-801.

Mager, U., and P. Nowak. 2012. 'Effects of student participation in decision making at school. A systematic review and synthesis of empirical research', *Educational Research Review*, 7: 38-61.

Marmot, M. 2018. 'Social determinants, capabilities and health inequalities: A response to Bhugra, Greco, Fennell and Venkatapuram', *Journal of Human Developement and Capabilities*, 19: 575-77.

Martin, R., and E.M. Murtagh. 2017. 'Effect of active lessons on physical activity, academic, and health outcomes: A systematic review', *Research Quarterly for Exercise and Sport*, 88: 149-68.

Maynard, B.R., K.T. McCrea, T.D. Pigott, and M.S. Kelly. 2012. 'Effects on school attendance among chronic truant students', *Campbell Systematic Reviews* 10: 1-83.

Moher, D., A. Liberati, J. Tetzlaff, and D.G. Altman. 2009. 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement', *PLoS Med*, 6: e1000097.

Moore, T.G., N. Arefadib, A. Deery, M. Keyes, and S. West. 2017. "The thousand days: An evidence paper." In. Parkville, Victoria: Centre for Community Child Health, Murdoch Children's Research Institute.

Muennig, P., L. Schweinhart, J. Montie, and M. Neidell. 2009. 'Effects of a prekindergarten educational intervention on adult health: 37-year follow-up results of a randomized controlled trial', *Research and Practice*, 99: 1431-37.

Mullender-Wijnsma, M.J., E. Hartman, J.W. de Greeff, S. Doolaard, R.J. Bosker, and C. Visscher. 2016. 'Physically active math and language lessons improve academic achievement: A cluster randomized controlled trial', *Pediatrics*, 137: 42-42.

National Health and Medical Research Council. 2009. "NHMRC levels of evidence and grades for recommendations for developers of guidelines " In.

Nigg, C.R., J.P. Allegrante, and M. Ory. 2002. 'Theory-comparison and multiple-behavior research: common themes advancing health behavior research', *Health Education Research*, 17: 670-79.

Nigg, C.R., and C.R. Long. 2012. 'A systematic review of single health behavior change interventions vs. multiple health behavior change interventions among older adults. ', *Translational Behavioral Medicine*, 2: 163-79.

NSW CESE. 2016. "How schools can improve literacy and numeracy and why it (still) matters." In.

NSW Department of Education. 2019. "School Excellence Framework evidence guide." In.

OECD. 2013. "PISA 2012 results in focus: What 15-year-olds know and what they can do with what they know." In.: OECD Publishing.

Ok, M.W., K. Rao, B.R. Bryant, and D. McDougall. 2017. 'Universal Design for Learning in Pre-K to Grade 12 classrooms: A systematic review of research', *Exceptionality*, 25: 116-38.

Owen, K.B., P.D. Parker, B. Van Zanden, F. MacMillan, T. Astell-Burt, and C. Lonsdale. 2016. 'Physical activity and school engagement in youth: A systematic review and meta-analysis', *Educational Psychologist*, 51: 129-45.

Patall, E.A., H. Cooper, and J.C. Robinson. 2008. 'Parent involvement in homework: A research synthesis', *Review of Educational Research*, 78: 1039-101.

Perry, L., and A. McConney. 2010. 'School socio-economic composition and student outcomes in Australia: Implications for educational policy', *American Journal of Education*, 54: 72-85.

Pfeiffer, B., A. Henry, S. Miller, and S. Witherell. 2008. 'Effectiveness of Disc 'O' Sit cushions on attention to task in second-grade students with attention difficulties', *American Journal of Occupational Therapy*, 62: 274-81.

Piasta, S.B., and R.K. Wagner. 2010. 'Developing early literacy skills: A meta-analysis of alphabet learning and instruction', *Reading Research Quarterly*, 45: 8-38.

Puzio, K., and G.T. Kolby. 2013. 'Cooperative learning and literacy: A meta-analytic review', *Journal of Research on Educational Effectiveness*, 6: 339.

Ramey, CT., FA. Campbell, M. Burchinal, ML. Skinner, DM. Gardner, and SL. Ramey. 2000. 'Persistent effects of early childhood education on high-risk children and their mothers', *Applied Developmental Science*, 4: 2-14.

Rhodes, C.M., L. Cordie, and M. Wooten. 2019. 'An examination of social capital among US Adults: Patterns that facilitate social well-being as measured by PIAAC', *International Journal of Learning, Teaching and Educational Research*, 18: 1-12.

Rohrbeck, C.A., M.D. Ginsburg-Block, J.W. Fantuzzo, and T.R. Miller. 2003. 'Peer-assisted learning interventions with elementary school students: A meta-analytic review', *Journal of Educational Psychology*, 95: 240-57.

Roorda, D.L., S. Jak, M. Zee, F.J. Oort, and H.M.Y. Koomen. 2017. 'Affective teacher-student relationships and students' engagement and achievement: A meta-analytic update and test of the mediating role of engagement', *School Psychology Review*, 46: 239-61.

Santangelo, T., and S. Graham. 2016. 'A comprehensive meta-analysis of handwriting instruction', *Educational Psychology Review*, 28: 225-65.

See, B.H., and S. Gorard. 2015. 'The role of parents in young people's education--a critical review of the causal evidence', *Oxford Review of Education*, 41: 346-66.

Semke, C.A., and S.M. Sheridan. 2012. 'Family-school connections in rural educational settings: A systematic review of the empirical literature', *School Community Journal*, 22: 21-47.

Sénéchal, M., and L. Young. 2008. 'The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review', *Review of Educational Research*, 78: 880-907.

Shernoff, E., and T. Kratochwill. 2007. 'Transporting an evidence-based classroom management program for preschoolers with disruptive behaviour problems to a school: an analysis of implementation, outcomes, and contextual variables', *School Psychology Quarterly*, 22 449–72. .

Shin, I.S., and J.Y. Chung. 2009. 'Class size and student achievement in the United States: A meta-analysis', *KEDI Journal of Educational Policy*, 6.

Sirin, S.R. . 2005. 'Socioeconomic status and academic achievement: A meta-analytic review of research', *Review of Educational Research*, 75: 417-53.

Sklad, M., R. Diekstra, M. De Ritter, J. Ben, and C. Gravesteijn. 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: 892-909.

Sklad, Marcin, Rene Diekstra, Monique De Ritter, Jehonathan Ben, and Carolien Gravesteijn. 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: 892-909.

Slavin, R.E., C. Lake, P. Hanley, and A. Thurston. 2012. "Effective programs for elementary science: A best-evidence synthesis." In.: Center for Research and Reform in Education.

Smith, Jean Louise M, Nancy J Nelson, Hank Fien, Keith Smolkowski, Derek Kosty, and Scott K Baker. 2016. 'Examining the efficacy of a multitiered intervention for at-risk readers in grade 1', *The Elementary School Journal*, 116: 549-73.

Sokolowski, A., Y. Li, and V. Willson. 2015. 'The effects of using exploratory computerized environments in grades 1 to 8 mathematics: A meta-analysis of research', *International Journal of STEM Education*, 2.

Spilt, J.L., H.M. Koomen, J.T. Thijs, and A. van der Leij. 2012. 'Supporting teachers' relationships with disruptive children: The potential of relationship-focused reflection', *Attachment & Human Development*, 14: 305-18.

Stockard, J., T.W. Wood, C. Coughlin, and C.R. Khoury. 2018. 'The effectiveness of direct instruction curricula: A meta-analysis of a half century of research', *Review of Educational Research*, 88: 479-507.

Suggate, S.P. 2016. 'A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions', *Journal of Learning Disabilities*, 49: 77-96.

Sung, Y.T., K.E. Chang, and T.C. Liu. 2016. 'The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis', *Computers & Education*, 94: 252-75.

Takacs, Z.K., E.K. Swart, and A.G. Bus. 2015. 'Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks: A meta-analysis', *Review of Educational Research*, 85: 698-739.

Thomas, T., K. Alexander, R. Jackson, and P.C. Abrami. 2013. 'The differential effects of interactive versus didactic pedagogy using computer-assisted instruction', *Journal of Educational Computing Research*, 49: 403-36.

Tingir, S., B. Cavlazoglu, O. Caliskan, O. Koklu, and S. Intepe-Tingir. 2017. 'Effects of mobile devices on K–12 students' achievement: a meta-analysis', *Journal of Computer Assisted Learning*, 33: 355-69.

Trentacosta, C.J., L.W. Hyde, D.S. Shaw, T.J. Dishion, F. Gardner, and M. Wilson. 2008. 'The relations among cumulative risk, parenting, and behavior problems during early childhood', *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 49: 1211-19.

Ttofi, M.M., and D.P. Farrington. 2011. 'Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review', *Journal of Experimental Criminology*, 7: 27-56.

Vandenbroucke, L., J. Spilt, K. Verschueren, C. Piccinin, and D. Baeyens. 2018. 'The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher–student interactions for children's executive functions', *Review of Educational Research*, 88: 125-64.

von Stumm, S. . 2017. 'Socioeconomic status amplified the achievement gap throughout compulsory education indepeedent of intelligence', *Intelligence*, 60: 57-62.

Walker, S.P., T.D. Wachs, S. Grantham, M.M. McGregor, C.A. Black, and S.L. Nelson. 2011. 'Inequality in early childhood: risk and protective factors for early child development', *The Lancet*, 378: 1325-38.

Wanzek, J., S. Vaughn, N. Scammacca, B. Gatlin, M.A. Walker, and P. Capin. 2016. 'Meta-Analyses of the Effects of Tier 2 Type Reading Interventions in Grades K-3', *Educational Psychology Review*, 28: 551-76.

Weaver, R.G., C.A. Webster, C. Egan, C.M.C. Campos, R.D. Michael, and S. Vazou. 2018. 'Partnerships for Active Children in Elementary Schools: Outcomes of a 2-Year Pilot Study to Increase Physical Activity During the School Day', *American Journal of Health Promotion*, 32: 621-30.

Whear, R., J. Thompson-Coon, K. Boddy, T. Ford, D. Racey, and K. Stein. 2013. 'The effect of teacher-led interventions on social and emotional behaviour in primary school children: A systematic review', *British Educational Research Journal*, 39: 383-420.

World Health Organisation. 2019. 'Global strategy on diet, physical activity, and health'. https://www.who.int/dietphysicalactivity/pa/en/.

Zeneli, M., A. Thurston, and C. Roseth. 2016. 'The influence of experimental design on the magnitude of the effect size -peer tutoring for elementary, middle and high school settings: A meta-analysis', *International Journal of Educational Research*, 76: 211-23.

Zenner, C., S. Herrnleben-Kurz, and H. Walach. 2014. 'Mindfulness-based interventions in schools-A systematic review and meta-analysis', *Frontiers in Psychology*, 5: 603.

Zoogman, S., S.B. Goldberg, W.T. Hoyt, and L. Miller. 2015. 'Mindfulness interventions with youth: A meta-analysis', *Mindfulness*, 6: 290-302.

APPENDICES

Appendix A: School improvement frameworks and tools

FISO (Victoria) **Priority areas and dimensions**	**SEF (New South Wales)** **Domains and elements**	**NSIT (National)** **Domains**	**OFSTED (United Kingdom)**	**NQS (ECEC Sector)** **Quality areas**
Excellence in teaching and learning: • Curriculum planning and assessment • Building practice excellence • Evidence-based high impact teaching strategies • Evaluating impact on learning Positive climate for learning: • Empowering students and building school pride • Setting expectations and promoting inclusion • Health and wellbeing • Intellectual engagement and self-awareness Community engagement in learning: • Building communities • Parents and carers as partners • Global citizenship • Networks with schools, services and agencies Professional leadership: • Building leadership teams • Instructional and shared leadership • Strategic resource management • Vision values and culture	Learning: • Learning culture • Wellbeing • Curriculum • Assessment • Reporting • Student performance measures Teaching: • Effective classroom practice • Data skills and use • Professional standards • Learning and development Leading: • Educational leadership • School planning, implementation and reporting • School resources • Management practices and processes	• Explicit improvement agenda • Analysis and discussion of data • Culture that promotes learning • Targeted use of school resources • Expert teaching team • Systematic curriculum delivery • Differentiated teaching and learning • Effective pedagogical practices • School-community partnerships	• Overall effectiveness • Quality of education • Behaviour and attitudes • Personal development • Leadership and management • Early years education • The sixth form	• Educational program and practice • Health and safety • Physical environment: • Staffing arrangements: • Relationships with children: • Collaborative partnerships with families and communities: • Governance and leadership:

Appendix B: Phase 1 search strategy and key terms

A1: Phase 1 search for systematic reviews & meta-analyses

1. Database: ERIC (EBSCOhost)
    a. (systematic review* or meta-analysis or DE Meta Analysis OR metaanalysis OR meta-synthesis OR metasynthesis OR meta-analytic OR metaanalytic)
    b. AND
    c. (early childhood or primary or elementary)
    d. Dates: 2012 up to 1st March 2018
    e. Results: 592
2. Database: Cochrane
    a. (school* or education* or classroom* or class or learning)
    b. AND
    c. (systematic review* or meta-analysis) AND (early childhood or primary or elementary)
    d. Dates: 2012 up to 9th August 2017
    e. Results: 219 (all excluded)
3. Database: Medline (Web of Science)
    a. (((TOPIC:(systematic review* OR meta-analysis) AND TOPIC: ((early) childhood OR primary) OR elementary)) AND TOPIC: ((((school*) OR education*) OR classroom*) OR class) OR learning)) AND TOPIC:(curriculum OR curricula))
    b. Dates: 2012 up to 17th August 2017
    c. Results: 63 (all excluded)
4. Database: A+ education
    a. systematic review or meta-analysis
    b. Dates: 2012 up to 17th August 2017
    c. Results: 132
5. Database: Psychinfo
    a. terms: class or classroom* or education* or learning or school*
    b. limit 10 to (("0830 systematic review" or 1200 meta analysis or 1300 metasynthesis) and childhood <birth to 12 years> and (100 childhood <birth to age 12 yrs> or 180 school age <age 6 to 12 yrs>) and yr="2012")
    c. Dates: 2012 up to 17th August 2017
    d. Results: 64
6. Database: ProQuest education database
    a. ab(meta-analysis OR "systematic review") AND ab("primary school" OR "primary level" OR "primary schools" OR elementary)
    b. Dates: 2012 up to 18th August 2017

Appendix C: Phase 2 search strategy and key terms

Note: for brevity, Ovid (PsycINFO) syntax is presented in the table below. Appropriate equivalent syntax was also developed and used in Ebsco (ERIC and ERC) and Cochrane database searches. Database-specific subject headings are collapsed for brevity (in practice they were entered for each database separately). Where database-specific terms were used they were combined with the domain topic keyword search using the 'OR' operator. In addition to searching ERIC, ERC, PsycINFO, and Cochrane, domain 11 searches were also conducted in Education Administration Abstracts.

Domain	Keywords	Date of search And limits
1	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (Phonics OR Phonemic-awareness OR phonological-awareness OR reading-comprehension OR reading-fluency OR repeated-reading OR ((Decode or decoding) ADJ3 (words OR skills)) OR vocabulary OR morphology OR explicit-instruction or direct-instruction or explicit-teaching or direct-teaching OR (math* ADJ3 (instruction OR teaching)) OR (science ADJ3 (education OR instruction OR teaching)) OR meta-cognitive-strateg*).ti,ab.	15 October 2018 (no date limit)
2	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (((differentiat* OR tailor* OR individuali*) ADJ3 (instruction OR feedback)) OR ((small-group) ADJ3 (teaching OR instruction OR tutoring)) OR progress-monitoring OR Tier-1 OR Tier-2 OR Tier-3 OR Tier-I OR Tier-II OR Tier-III).ti,ab. Data-base specific terms: individualized instruction, small group instruction, progress monitoring, response to intervention, tutoring, remedial instruction, remedial reading, remedial mathematics	16 October 2018 (no date limit)
3	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (Reciprocal-learning OR reciprocal-teaching OR cooperative-learning OR peer-teaching or cross-age-teaching or PALS OR peer-assisted-learning-strategies OR CIRC OR CORI OR Cooperative-Integrated-Reading-and-Composition OR Conceptual-Oriented-Reading-Instruction).ti,ab. Data-base specific terms: cooperative learning, peer teaching, reciprocal teaching, cross age teaching, peer Tutoring, collaborative learning	26 October 2018 (no date limit)
4	(Early-Childhood-Education OR grade-1 OR grade-2 OR grade-3 OR grade-one OR grade-two OR grade-three OR kindergarten OR first-grade OR second-grade OR third-grade OR reception OR Key-Stage-1 OR Key-Stage-One).ti,ab. AND (Primary-school* OR elementary-school* OR primary-education or elementary-education OR classroom* OR student* OR school* OR general-education).ti,ab. AND (Physical-activit* OR exercise* OR perceptual-motor-learning OR perceptual-motor-coordination OR psychomotor-skills OR perceptual-motor-program).ti,ab. AND (achievement OR academic OR read* OR writ* OR spell* OR math* OR science OR cognit* OR educational-attainment OR achievement-gap).ti,ab. AND	Limited to 2008-2018

Domain	Keywords	Date of search And limits
	(RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab.	
5	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (iPad OR tablet OR hand-held-device* OR handheld OR smartphone OR PDA OR interactive-whiteboard OR computer-assisted-instruction OR CAI OR game-based-learning OR ((computer AND game*) ADJ4 (teaching OR learning)) OR educational-technology OR educational-equipment OR educational-media OR electronic-learning OR computer-managed-instruction).ti,ab. Database-specific terms: technology uses in education, educational technology, educational equipment, educational media, electronic learning, computer assisted instruction, computer managed instruction, integrated learning systems, teaching aids, computers in education, electronic classrooms, multimedia systems in education, media programs (education), audiovisual education, programmed instruction, educational audiovisual aids, teaching machines, intelligent tutoring systems, learning management systems, mobile devices	16 October 2018 (no date limit)
6	(Early-Childhood-Education OR grade-1 OR grade-2 OR grade-3 OR grade-one OR grade-two OR grade-three OR kindergarten OR first-grade OR second-grade OR third-grade OR reception OR Key-Stage-1 OR Key-Stage-One).ti,ab. AND (Primary-school* OR elementary-school* OR primary-education or elementary-education OR classroom* OR student* OR school* OR general-education).ti,ab. AND (Natural-light* OR artificial-light* OR green-space* OR seating OR reading-corner OR graphic-reminder* OR noise OR heating OR cooling OR classroom-setting* OR physical-environment OR playground* OR play-ground OR outdoor-area* OR outdoor-space* OR outdoor OR play-space* OR playspace OR physical-space* OR classroom-decoration* OR classroom-design OR classroom-display* OR school-setting OR space-utili#ation OR school-building* OR school-facilit* OR school-premises OR literacy-rich-environment OR natural-environment OR plant* OR grow-plants OR trees OR flower* OR garden*).ti,ab. AND (RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. Database-specific terms: school buildings, educational facilities design, space utilization, playgrounds, human factors engineering, physical environment, climate control, interior design, interior space, classroom design, school space, school environment, school decoration, school lighting, classroom design & construction, classroom environment, play environments, play equipment, noise (work environment), heating & ventilation of school buildings, air conditioning in school buildings, playgrounds, school facilities, nature (environment), environmental psychology, physical comfort, furniture	16 October 2018 Limit to 2008-2018
7a	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (Teacher-student-ratio OR teacher-distribution OR instructor-student-ratio OR instructor-distribution OR faculty-student-ratio OR pupil-teacher-ratio OR teacher-pupil-ratio OR student-faculty-ratio OR student-teacher-ratio).ti,ab.	17 October 2018 (no date limit)

Domain	Keywords	Date of search And limits
	Database-specific terms: Small classes, teacher student ratio, class size, teacher distribution	
7b	(RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. AND (Teacher-student-ratio OR teacher-distribution OR instructor-student-ratio OR instructor-distribution OR faculty-student-ratio OR pupil-teacher-ratio OR teacher-pupil-ratio OR student-faculty-ratio OR student-teacher-ratio).ti,ab.	21 November 2018 Limit to 2008-2018
8	((Early-Childhood-Education OR grade-1 OR grade-2 OR grade-3 OR grade-one OR grade-two OR grade-three OR kindergarten OR first-grade OR second-grade OR third-grade OR reception OR Key-Stage-1 OR Key-Stage-One).ti,ab. OR ((Primary-school* OR elementary-school* OR primary-education or elementary-education OR classroom* OR student* OR school* OR general-education).ti,ab.) AND (Student-leadership OR student-empowerment OR student-organi#ation* OR student-council* OR student-representative-council* OR student-representative-committee OR student-representative* OR SRC OR student-voice* OR student-govern* OR ((leadership-training) ADJ4 (student*)).ti,ab. AND (RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. Database-specific terms: student government, student empowerment, student leadership, student responsibility, leadership training, student participation in administration, self-efficacy in students	21 November 2018 Limit to 2008-2018
9	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND ((Peer-management OR peer-counselling OR peer-mediation OR peer-support OR buddy-program*) OR ((differen* OR divers*) ADJ4 (tolera* OR accept* OR celebrat*)) OR ((cultural-diversity) ADJ4 (tolera* OR accept* OR celebrat*)) OR ((belong* OR connect*) ADJ3 (sense-of OR feel*)) OR (mindfulness) OR ((gratitude OR grateful* OR empath*) ADJ4 (teach* OR strateg* OR practice)) OR chaplain* OR school-counselling OR school-counsellor).ti,ab. **Database-specific terms:** Peer counselling, peer mediation, student diversity, cultural differences, sense of community, peer relations	17 October 2018 (no date limit)
10a	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (((Positive-relationship*) ADJ6 (teacher* OR student*)) OR ((quality-of-relationship*) ADJ6 (teacher* OR student*)) OR ((relationship-quality) ADJ6 (teacher* OR student*)) OR (teacher-student-relationship* OR teacher-child-relationship* OR teacher-student-interaction* OR student-teacher-relationship* OR child-teacher-relationship OR student-teacher-interaction)).ti,ab. **Database-specific terms:** teacher student relationship, teacher expectations, teacher expectations of students, student school relationship, classroom environment, classroom communication	17 October 2018 (no date limit)
10b	(RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. AND	27 November 2018

Domain	Keywords	Date of search And limits
	(((Positive-relationship*) ADJ6 (teacher* OR student*)) OR ((quality-of-relationship*) ADJ6 (teacher* OR student*)) OR ((relationship-quality) ADJ6 (teacher* OR student*)) OR (teacher-student-relationship* OR teacher-child-relationship* OR teacher-student-interaction* OR student-teacher-relationship* OR child-teacher-relationship OR student-teacher-interaction)).ti,ab.	Limit to 2008-2018
11	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (School-leader* OR principal* OR school-turn-around OR ((turnaround) ADJ5 (school*)) OR leadership-training OR leader-qualit* OR leadership-qualit* OR leadership-effectiveness OR effective-leadership OR effective-leader* OR transformational-leadership OR school-management OR school-administration OR school-improvement OR school-vision OR school-performance).ti,ab. **Database-specific terms:** training of school principals, school turnaround, leadership, leadership qualities, leadership effectiveness, leadership training, transformational leadership, professional development, staff development, educational leadership, teacher leadership, school administration, school administrators, career development, continuing education, professional education, school principals, educational administration	17 October 2018 (no date limit)
11b	(RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. AND (Principal-leadership OR principal-professional-development).ti,ab.	12 December 2018 Limit to 2008-2018
12	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (School-family-partnership* OR family-school-partnership* OR ((school*) ADJ6 (famil* AND collaborat*)) OR ((school*) ADJ6 (partner* AND famil*)) OR ((parent AND teacher) ADJ6 (relationship* OR partner* OR collaborat*)) OR ((parent*) ADJ3 (involve* OR participat*)) OR ((family) ADJ3 (involve* OR participat*)) OR school-family OR family-school OR school-home OR school-parent OR parent-school).ti,ab. Database-specific terms: family school relationship, parent school relationship, partnerships in education, school community relationship, school involvement, parent teacher conferences, parent teacher cooperation, parent-teacher relationships, family involvement, community & school, parent participation in education, parent-administrator relationships, parental involvement	22 October 2018 (no date limit)
13a	(Meta-analysis OR Meta-analytic OR metaanaly*).ti,ab. AND (School-community-partnership* OR community-school-partnership* OR (school* ADJ6 (community* AND collaborat*)) OR (school* ADJ6 (partner* AND community*)) OR ((community AND teacher) ADJ6 (relationship* OR partner* OR collaborat*)) OR (community* ADJ3 (involve* OR participat*)) OR (community ADJ3 (involve* OR participat*)) OR school-community OR community-school OR (school ADJ6 (external organi#ation* OR agency OR agencies OR local-business*)) OR (school ADJ6 (co-location OR integration-of-services OR integrate*)) OR (local ADJ6 agenc*) OR (local ADJ6 service*) OR (local ADJ6 organisation*) OR (local ADJ6 network*) OR school-university OR university-school).ti,ab.	22 October 2018 (no date limit)

Domain	Keywords	Date of search And limits
	Database-specific terms: partnerships in education, school community relationship, school community programs, school business relationship, school Involvement, community & school, school-state government relationships	
13b	(RCT OR randomi#ed-control* OR ((random*) ADJ2 (assign* OR allocat*)) OR randomly OR Randomi#e* OR clinical-trial* OR Experimen* OR Cluster-RCT OR CRCT OR Cluster-Random).ti,ab. AND (School-community-partnership* OR community-school-partnership* OR (school* ADJ6 (community* AND collaborat*)) OR (school* ADJ6 (partner* AND community*)) OR ((community AND teacher) ADJ6 (relationship* OR partner* OR collaborat*)) OR (community* ADJ3 (involve* OR participat*)) OR (community ADJ3 (involve* OR participat*)) OR school-community OR community-school OR (school ADJ6 (external-organi#ation* OR agency OR agencies OR local-business*)) OR (school ADJ6 (co-location OR integration-of-services OR integrate*)) OR (local ADJ6 agenc*) OR (local ADJ6 service*) OR (local ADJ6 organisation*) OR (local ADJ6 network*) OR school-university OR university-school).ti,ab. AND (Early-Childhood-Education OR grade-1 OR grade-2 OR grade-3 OR grade-one OR grade-two OR grade-three OR kindergarten OR first-grade OR second-grade OR third-grade OR reception OR Key-Stage-1 OR Key-Stage-One OR Primary-school* OR elementary-school* OR primary-education or elementary-education OR primary-student* OR elementary-student*).ti,ab. Database-specific terms: partnerships in education, school community relationship, school community programs, school business relationship, school involvement, community & school, school-state government relationships	4 December 2018 Limit to 2008-2018

Appendix D: PRISMA 2009 checklist

Section/topic	#	Checklist item
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.

Section/topic	#	Checklist item
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Appendix E: NICE Quality and Bias Checklist

Paper: ref #_____

1.1 Is the source population or source area well described?
1.2 Is the eligible population or area representative of the source population or area?
1.3 Do the selected participants or areas represent the eligible population or area?

2.1 Allocation to intervention (or comparison). How was selection bias minimised?
2.2 Were interventions (and comparisons) well described and appropriate?
2.3 Was the allocation concealed?
2.4 Were participants or investigators blind to exposure and comparison?
2.5 Was the exposure to the intervention and comparison adequate?
2.6 Was contamination acceptably low?
2.7 Were other interventions similar in both groups?
2.8 Were all participants accounted for at study conclusion?

3.1 Were outcome measures reliable?
3.2 Were all outcome measurements complete?
3.3 Were all important outcomes assessed?
3.4 Were outcomes relevant?
3.5 Were there similar follow-up times in exposure and comparison groups?
3.6 Was follow-up time meaningful?

4.1 Were exposure and comparison groups similar at baseline? If not, were these adjusted?
4.2 Was intention to treat (ITT) analysis conducted?
4.3 Was the study sufficiently powered to detect an intervention effect (if one exists)?
4.4 Were the estimates of effect size given or calculable?
4.5 Were the analytical methods appropriate?
4.6 Was the precision of intervention effects given or calculable? Were they meaningful?

5.1 Are the study results internally valid (i.e. unbiased)?
5.2 Are the findings generalizable to the source population (externally valid)?

Overall rating
++ ALL or most of the criteria have been fulfilled (75%)
+ SOME of the criteria have been fulfilled (51 – 74%)
-FEW or NO checklist criteria have been fulfilled (50% and below)

Appendix F: Strategy & domain-level evidence ranking systems

OVERALL RANKING OF THE EVIDENCE-STUDY LEVEL	
	Definition
Well Supported	Clear, consistent evidence of benefit. No evidence of harm or risk to participants. A well conducted[24] systematic review or meta-analysis found the intervention to be more effective than a control group on at least one child valid outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). Populations examined are similar to, and results are sensible to apply to, the Australian primary[25] school context.
Supported	Evidence suggestive of benefit but more evidence needed. No evidence of harm or risk to participants. A systematic review or meta-analysis of moderate quality[26] found the intervention to be more effective than a control group on at least one child valid outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the review are sensible to apply to primary school age students. Populations examined may be somewhat different to the Australian population; affecting generalisability to the Australian context.
Promising	No evidence of harm or risk to participants. At least one RCT with low to moderate risk of bias found the intervention to be more effective than a control group on at least one valid child outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the study are sensible to apply to primary school age children[27], though populations may be somewhat different to the Australian population.
Preliminary	No evidence of harm or risk to participants. At least one QES with low risk of bias found the intervention to be more effective than a control group on at least one valid child outcome measure (i.e. cognition, language, academic achievement, social-emotional functioning). The results of the study are sensible to apply to primary school age children[28], though populations may be somewhat different to the Australian population.
Not Supported	A well conducted systematic review or meta-analysis or at least one RCT found the intervention to be ineffective across several primary outcomes

[24] To be considered well-conducted, meta-analyses and systematic reviews had to receive a PRISMA rating indicating low risk of bias (++) and at least 50% of included studies had to be RCTs, QESs, or matched comparison designs.
[25] For meta-analyses and systematic reviews to be considered relevant to the early years of school, at least 50% of included studies had to involve elementary school students or results reported separately for elementary students.
[26] Moderate quality means the meta-analysis or review received a PRISMA rating indicating moderate risk of bias (+) and included at least 50% RCT, QES, or matched-comparison designs.
[27] At least 50% of participants, or the average age of participants, must be within the primary school range (i.e. 4 years to 12 years).
[28] At least 50% of participants, or the average age of participants, must be within the primary school range (i.e. 4 years to 12 years).

OVERALL RANKING OF THE EVIDENCE-STUDY LEVEL	
	compared with a control group. The overall weight of the evidence does not support the benefit of the practice.
Concerning Practice	A well conducted systematic review or meta-analysis reported that the direction of effects was undesirable across several outcomes. At least 1 RCT with low risk of bias showed the practice to have a negative effect.
Unknown	The intervention has not been adequately assessed. Available meta-analyses, reviews, or RCTs are limited either in terms of quality (low PRISMA/NICE rating) or relevance (to primary school age population).

OVERALL RANKING OF THE EVIDENCE-DOMAIN LEVEL	
	Definition
Well Supported	At least two meta-analyses/systematic reviews identified different strategies rated Well Supported
Supported	At least one meta-analysis/systematic review identified a strategy rated Supported or Well Supported
Promising	At least two high quality RCTs identified different types of strategies with demonstrated effectiveness
Preliminary	At least one high quality quasi-experimental study or moderate quality RCT identified an effective strategy in this domain
Mixed	There are conflicting findings for similar strategies identified in equal quality studies (e.g. one high-quality meta-analysis suggests the strategy is not effective, while another high-quality meta-analysis suggests it is supported)
Not Supported	The strategies identified in this domain were consistently rated Not Supported or Concerning practices
Unknown	No relevant meta-analyses, systematic reviews, RCTs, or quasi-experimental studies were identified in this domain OR the evidence for strategies identified in the domain were rated Unknown due to poor methodological quality or low relevance to primary school age children

Appendix G: Phase 2 search Prisma flow diagram

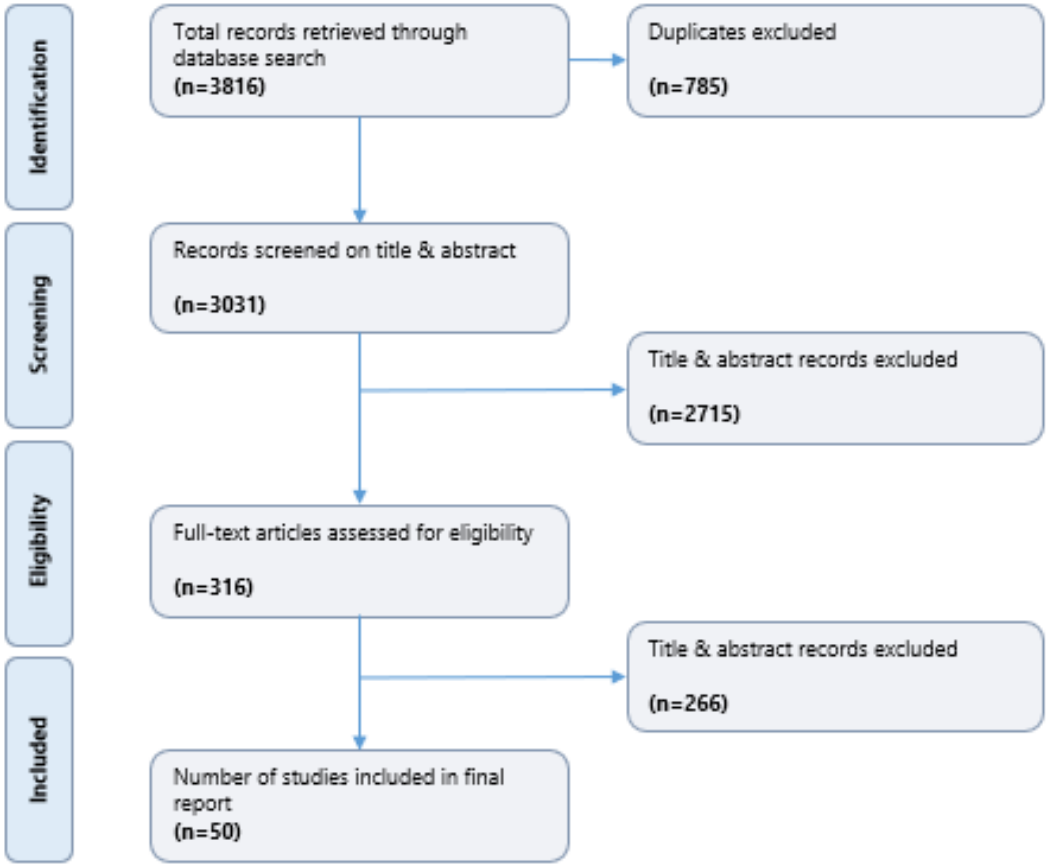

Figure 2: Prisma flow chart of phase 2 searches

Appendix H: Phase 2 search records screened by domain

Domain (MA/RCT-level)	N after removing duplicates, books, theses, reports, case-studies etc.	N coded include at Abstract	N include at Full text (excluding duplicates with Phase 1)
1: Curriculum content and delivery (MA)	259	64	7
2: Differentiated teaching strategies (MA)	126	22	3
3: Peer teaching and co-operative learning (MA)	106	19	5
4: Use of physical activity for academic outcomes (RCT)	63	13	5
5: Technology-assisted teaching (MA)	360	80	5
6: Physical environment to optimise learning (RCT)	147	5	3
7: Class size and teacher-student ratios (MA)	27	8	0
7b: RCT-level class size search (2009-Nov 2018)	78	12	0
8: Student empowerment and leadership (RCT)	314	3	0
9: SEB specific strategies (mindfulness, peer mediation, chaplaincy etc) (MA)	361	8	3
9b: physical activity on SEB outcomes only (RCT)	152	6	0
10: Staff-student relationships (MA) (correlational ok)	96	12	5
10b: RCTs of interventions to improve staff-student relationships	161	27	5
11a: Leadership development (MA)	262	32	0
11b: Leadership development (RCT)	83	7	1
11b: supplementary search	30	0	0
12: Partnerships with families (MA)	177	21	3
13: Community-School Partnerships (MA)	61	4	1
13b: Community-School Partnerships (RCT)	168	15	4

Appendix I: List of publications included in review by search phase and domain

Author (date)	Study ID #	Study type (SR/MA/RCT)	Search Phase (1 or 2)	Content relevant to domain
Abrami et al. (2015)	#58	MA	1	Curriculum
Abrami & Russo (2015)	#958	MA	2	Technology
Abry et al (2013)	#596	RCT	2	Teacher-student relationships
Amlani et al (2016)	#305	REE	2	Physical environment
Barbero (2012)	#109	SR	1	Social Emotional & Behavioural
Blank & Alas (2009)	#90	MA	1	Staff and leadership development
Cappella et al (2012)	#598	RCT	2	Teacher-student relationships
Carbonneau et al.	#88	MA	1	Curriculum
Carsley et al (2018)	#328	MA	2	Social Emotional & Behavioural
Celio et al (2011)	#602	MA	2	Community-School Partnerships
Cheung & Slavin (2012)	#72	MA	1	Technology
Cornelius-White (2007)	#342	MA	2	Teacher-student relationships
Davies (2013)	#91	SR	1	Community-school partnerships
Deunk et al (2018)	#350	MA	2	Differentiated instruction
Dietrichson et al. (2017)	#94	MA	1	Peer teaching; Technology; SEB; Staff & Leadership development
Donnelly et al (2009)	#352	RCT	2	Physical activity
Donnelly et al (2017)	#353	RCT	2	Physical activity
Dunst et al (2015)	#153	MS	1	Staff & Leadership development;
Durlack (2011)	#959	MA	2	SEB
Elbaum et al (1999)	#361	MA	2	Peer teaching
Elleman (2017)	#150	MA	1	Curriculum
Erion (2006)	#366	MA	2	Family-School partnerships
Fernandez et al (2015)	#599	RCT	2	Teacher-student relationships
Fisher et al (2014)	#368	RME	2	Physical environment
Glass & Smith (1979)	#100	MA	1	Class size
Goodwin et al (2013)	#383	MA	2	Curriculum
Gordon et al (2015)	#384	MA	2	Curriculum
Grahan & Santangelo (2014)	#68	MA	1	Curriculum
Graham et al (2015)	#74	MA	1	Differentiated instruction
Graham et al (2018)	#387	MA	2	Curriculum
Gunter & Shao (2016)	#3	MA	1	Physical environment
Hammill et al (2006)[29]	#393	MA	2	Curriculum
Harvey et al (2018)	#396	RCT	2	Physical activity
Henry et al (2017)	#604	QED	2	Community-school partnerships

[29] reanalysis of Ehri 2001

Author (date)	Study ID #	Study type (SR/MA/RCT)	Search Phase (1 or 2)	Content relevant to domain
Hoglund et al (2012)	#612	QED	2	Community-school partnerships
Jacob et al (2015)	#603	RCT	2	Staff and leadership development
Jeynes (2012)	#75	MA	1	Family-school partnerships
Kim & Hill (2015)	#53	MA	1	Family-school partnerships
Korpershoek et al (2016)	#19	MA	1 & 2	Teacher-student relationships & SEB
Kraft et al (2018)	#427	MA	2	Staff and leadership development
Kunsch et al (2007)	#430	MA	2	Peer teaching
Lee et al (2013)	#440	MA	2	Technology
Lee & Stuart (2013)	#615	QED	2	Community-school partnerships
Lei et al (2016)	#441	MA	2	Teacher-student relationships
Leithwood & Sun (2012)	#73	MA	1	Staff and leadership development
Leung (2015)	#443	MA	2	Peer teaching
Li et al (2014)	#444	RCT	2	Physical activity
Lou et al (2000)	#454	MA	2	Differentiated instruction
Ma et al (2016)	#2	MA	1	Family-school partnership
Magerr & Nowak (2011)	#93	SR	1	Student empowerment
Martin & Murtagh (2017)	#48	SR	1	Physical activity
Maynard et al (2012)	#166	MA	1	Social Emotional & Behavioural
Mullender-Wijnsma et al (2016)	#477	RCT	2	Physical activity
Ok et al. (2016)	#23	SR	1	Differentiated instruction; Technology
Owen (2016)	#34	SR/MA	1	Physical activity
Patall et al (2008)	#486	MA	2	Family-school partnerships
Pfeiffer et al (2008)	#487	RCT	2	Physical environment
Piasta et al (2010)	#490	MA	2	Curriculum; Differentiated instruction
Puzio (2013)	#110	MA	1	Peer teaching
Rohrbeck et al (2003)	#497	MA	2	Peer teaching
Roorda et al (2017)	#498	MA	2	Teacher-student relationships
Santangelo & Graham (2016)	#46	MA	1	Curriculum; Differentiated instruction; Technology
See & Gorard (2015)	#116	SR	1	Family-school partnership
Semke & Sheridan	#87	SR	1	Family-school partnership
Sénéchal & Young (2008)	#513	MA	2	Family-school partnership
Shin & Chung (2009)	#112	MA	1	Class size
Sklad et al (2012)	#81	MA	1	Social Emotional & Behavioural
Slavin et al (2012)	#55	SR	1	Curriculum; Technology; Staff & leadership development
Spilt et al (2012)	#600	RCT	2	Teacher-student relationship

Author (date)	Study ID #	Study type (SR/MA/RCT)	Search Phase (1 or 2)	Content relevant to domain
Sokolowski et al (2015)	#516	MA	2	Technology
Stockard et al (2018)	#521	MA	2	Curriculum
Suggate (2016)	#523	MA	2	Curriculum
Sung et al (2016)	#528	MA	2	Technology
Takacs et al (2015)	#532	MA	2	Technology
Tfioti (2011)	#618	MA	2	Social Emotional & Behavioural; Family-School Partnerships
Thomas et al. (2013)	#76	MA	1	Technology
Tingir et al. (2017)	#149	MA	1	Technology
Vandenbroucke et al (2018)	#552	MA	2	Teacher-student relationship
Weaver et al (2018)	#617	QED	2	Community-school partnerships
Whear et al (2013)	#51	SR	1	Social Emotional & Behavioural; Staff & Leadership Development
Zeneli et al (2016)	#577	MA	2	Peer teaching
Zenner et al (2014)	#578	MA	2	Social Emotional & Behavioural
Zoogman et al (2015)	#582	MA	2	Social Emotional & Behavioural

MA: Meta-analysis; MS: Meta-synthesis; QED: Quasi-Experimental Design; RCT: Randomised Controlled Trial; REE: Repeated Exposure Experiment; SR: Systematic Review;

Appendix J: Details of included studies by domain

Domain 1: Application of pedagogical content knowledge

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
#58 Abrami et al. (2015)	• Meta-analysis • 684 studies (341 relevant ES) • RCT, QES, & PP • Published and unpublished • Search range: Up to 2009	• 6 years + (including tertiary and non-school settings) • Countries: NR[c] • Settings: diverse • Sample sizes: NR	• Various instructional interventions to improve critical thinking • Control: various (non-exposure) • Duration: at least 3 hours in total • Population Focus: Universal	• Critical thinking (standardised) • Achievement & content specific measures (teacher or researcher created)	• Overall g+ = 0.39*; p<0.001 , #ES=867, • Students 6-10yrs: g+ = 0.37*, p<0.05, #ES=49 Moderators assessed: age, subject, dose, instructor training. Follow up: NR	++ (78%)	Pr: 14% EYS: 9% ES for age 6-10 years reported	Supported
#88 Carbonneau et al. (2012)	• Meta-analysis • 55 studies • 13 Experiments, 30 QES, 12 PP • Published & unpublished • Search range: 1955-2010	• K-Post - secondary • Countries: NR • Settings: NR • n=7237	• Instruction utilising manipulatives to teach mathematics • Control: No manipulatives nor iconic representation, but same math concepts taught • Duration: Mean 25, range 1-180 days • Population Focus: NR	• Academic (mathematics: retention, problem-solving, transfer)	• Small to moderate positive effects across outcomes • Overall g= 0.37*, p≤0.001(k=55) • Overall ES for children 3-6 years (ES=0.33*, p<0.01, k=10) • Larger ES with high instructional guidance for some measures (g=0.46* vs 0.29*,p=0.01) • Follow Up: NR	+ (70%)	Pr: 75% EYS: 42% But results for 3-6 presented separately.	Supported
#150 Elleman (2017)	• Meta-analysis • 25 studies • RCTs & QES	• K-12 • Countries: NR • Settings: NR	• Inference instruction • Control: usual practice	• Comprehension (literal & inferential)	Comprehension effects: • General, d=0.58*, p≤0.01, k=13.	++ (84%)	Pr: 82% EYS: 27% (based on 33 cohorts)	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
	• Published & unpublished • Search range: 1950-2014	• n= 1752	• Duration: 70% <10hours, range: 2-67 hours • Population Focus: Universal and targeted		• Inferential, d=0.68*, p≤0.04, k=25 • Literal, d=0.28*, p≤0.04, k=18 [d=0.97*, p<0.05 less skilled, d=0.06, ns skilled readers]. Follow up: NR			
#383 Goodwin & Ahn (2013)	• Meta-analysis • 30 studies (92 effects) • Experimental & QED • Published & unpublished • Search range: 1980+ (last 2010)	• PreK-9 • Countries: NR • Settings: NR • n= 7042	• Morphological interventions • Control: usual practice, no treatment, alternative • Duration: 0.5 -67 hours • Population Focus: NR	• Literacy outcomes	Literacy outcomes: • Overall d= 0.32*, p<0.05, k=92 • Decoding d=0.59*, p<0.05, k=13; Phonological awareness d=0.48*, p<0.05, k=11; Morphological knowledge d =0.44*, p<0.05, k=11; Vocabulary d =0.34*, p<0.05, k=9 ; Spelling d = 0.30*, p<0.05, k=23; Reading comprehension d=0.09, ns, k =14; fluency d=–0.05, ns, k=11 Early elementary: • d =0.68*, p<0.05, k=17 Follow up: 2 wks-8months	+ (74%)	Pr: 90% EYS: 20%	Supported
#384 Gordon et al (2015)	• Meta-analysis • 13 studies • Experimental & QED • Published • Search range: Up to 2014 (earliest study 2004)	• Ages 4-9 (K-4) • Countries: NR • Settings: NR • n=901	• Music education • Control: no treatment, less intensive & alternative • Duration: 2-90 hours • Population Focus: NR	• Phonological awareness • Reading fluency	• Overall, d = 0.20; p =0.01 ,k=13 • Phonological Awareness d=0.20*,p =0 .01, k=18, • Rhyming d= 0.18, ns, k=7, • Reading Fluency d= 0.16, ns, k=5 Follow Up: NR	++ (83%)	Pr: 100 % EYS: 77%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
#387 Graham et al (2018)	• Meta-analysis • 47 studies • Experimental (9) & QED (38) • Published and unpublished • Search range: Up to 2017 (earliest study 1984)	• Grades preK-12 • Countries: 89% US, 2 Canada, 1 Iran, 1 Turkey • Settings: Urban, suburban & rural • n=30520	• Balanced reading and writing instructions • Control: NR • Duration: 2weeks-8 years • Population Focus: Universal and targeted	• Reading (comprehension, decoding, vocabulary) • writing (quality, mechanics, output)	Overall reading: small to moderate effect, g= 0.33*, p<0.001, k=38) • comprehension (g = 0.39*, p<0.001; k=23); decoding (g=0.53*, p <0.001, k=12); vocabulary (g = 0.35*, p= 0.002; k=9). Overall writing: small to moderate effect, g = 0.37*, p<0.001, k=37 • writing quality (g=0.47*,p=0.001, k=22); mechanics (g=0.18*, p<0.001, k=22); output (g =0.69*, p<0.001, k=22). EYS (preschool to grade 1): • reading, g= 0.46*, p<0.001, (k=7); writing performance, g=0.33*, p=0.18, (k=6) Follow up: NR	++ (85%)	Pr:66% EYS: 45%	Well Supported
#68 Graham & Santangelo (2014)	• Meta-analysis • 53 studies • RCT & QES • Published & unpublished • Search range: Up to Oct 2012	• K-12 • Countries: NR (authors US-based) • Settings: NR • n~6037 (I: 3244)	• Direct teaching of spelling skills • Control: no or less formal spelling instruction • Duration: NR • Population Focus: Universal (some studies target by literacy level)	• Spelling • Phonological awareness • Reading	• Improved spelling compared to no/unrelated instruction (d = 0.54*, p<0.001, k=25 overall, and d=0.63*, p<0.001 for K-2) or informal/incidental approaches (overall d = 0.43*, p<0.001, k=23). • Improvements in phonological awareness (d = 0.51*, p<0.05; k=7)and reading skills (d=0.44*, p<0.001; k=20) • GradeK-2 spelling: d=0.63*, k=2 Follow up: 6 studies 1wk to 6mth maintenance	++ (89%)	Pr: 85% EYS: 50% (based on effect sizes)	Supported
#393 Hammill et al (2006)	• Re-analysis of previous	• K-6 • Countries: NR	• Phonics instruction • Control: NR	• Literacy: decoding, reading,	Overall, significant (p<.05) positive effects:	++(76%) (based on Ehri)	Pr: 100% EYS: 48%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
	meta-analysis • 38 studies (66 effects) • Experimental & QED • Published • Search range: From 1970-2000	• Settings: NR • Sample size: NR	• Duration: NR • Population Focus: NR	spelling, comprehension	• decoding regular words (d=0.67, #ES=30), decoding pseudowords (d=0.60, #ES = 40), reading miscellaneous words (d=0.40, #ES =59), spelling words (d=0.35, #ES =7), reading text orally (d=0.25, #ES =16) comprehending text (d=0.27, #ES =35). Early years (all significant (p<.05) positive effects) • Overall ES for kindergarten and first grade (d = .55*, #ES = 30) • Overall ES for kindergarten (d = .56, #ES =7) and first grade (d =.54*, #ES = 23). • K-1: Decoding pseudowords d=0.67*, #ES 14; Reading miscellaneous words d=0.45*, #ES=23; Reading text only d=0.23*, #ES=6; Comprehending text d=0.51* #ES=11; decoding skills d=0.83*, #ES=8 Follow up: NR (NR in Ehri paper)			
Piasta et al 2010	• Meta-analysis • 63 studies (3-11 rel.) • 39 RCT/CRCT,21 QED, 3 include both • Published & Unpublished • Search range NR, dates of	• PK-6 • Countries: Canada, France, Germany, Israel, the Netherlands, New Zealand, the UK, USA • Setting: NR • n=8,469	• Group size as moderator for literacy instruction • Control: NR • Duration: 120-5,793 mins • Population Focus: Universal	• Literacy skills: Letter name , Letter sound, Letter writing, Letter name fluency	Small group vs whole class: Letter name: 0.52*, k=21 vs. 0.24*, k=4 Letter sound: 0.73, k=2 vs. 0.48, k=11 Letter writing 0.56, k=2 vs 0.60, k=3 Letter name fluency: 0.07, k=7 vs 0.06, k=5	++(78%)	Pr:100% EYS: 92%	Well Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
	included papers 1980-2007							
#46 Santangelo & Graham (2016)	• Meta-analysis • 76 studies (80 experiments) • RCTs & QES • Published & Unpublished • Search range: 1931-2015	• K-12 • Countries: NR (but searched US and Australian journals) • Settings: NR • Sample sizes: NR	• Handwriting (HW) instruction methods e.g. motor instruction, individualised HW instruction, technology, HW programs. • Control: non-handwriting instruction • Duration: NR • Population Focus: Universal	• Handwriting legibility and fluency	Positive effect of HW instruction (vs. no HW instruction) • legibility g=0.59*, k=20), • fluency g=0.63*, k=15 Effects on legibility: • Self-evaluation (g=0.66, ns, k=4,) • Teaching individual letters with motion (g=0.26, ns, k=5) • Copying letters from models (g= 0.26, ns, k=4) Follow up: NR	++ (87%)	Pr: 93% (overall), 84% (domain) EYS: 73% (overall), 52% (domain)	Supported
#55 Slavin et al (2012)	• Systematic review • 17 studies • RCTs, QES, & matched controls • Published and unpublished • Search range: 1980-2011	• K-6 • Countries: England, USA, Taiwan, Kuwait • Settings: Urban & Rural • Sample sizes: NR	• Programs and practices used in elementary science • Control: alternative program or standard methods • Duration: 4 weeks minimum • Population Focus: NR	• Academic (science specific)	• Inquiry-based programs using science kits showed no effect (d'= +0.02, k=4) Follow up: NR	+ (52%)	Pr: 100% EYS: 6%	Not Supported
#521 Stockard et al (2018)	• Meta-analysis • 328 studies • Mixed methods	• K-4 • Countries: USA and others not specified	• Direct instruction • Control: NR • Duration: few days to 6 years • Population Focus: Universal and targeted	• Reading • Mathematics • Spelling • Language	Overall (all p<0.001): • Reading d= 0.51*; k=226. • Maths d=0.55*, k= 70 • Language d=0.54*, k=56 • Spelling d= 0.66*; k =52 Early Years (kinder):	-(48%)	Pr: 100% EYS:NR	Unknown

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings/Conclusions	Quality Rating	Overall Relevance[b]	Evidence Rating
	• Published and unpublished • Search range: 1966-2016	• Settings: about half urban • Sample size: NR			• Maths d=0.17, p<0.05; • spelling d=0.37 p<0.01. • Reading/spelling: NR Follow up: NR			
#523 Suggate (2016)	• Meta-analysis • 16 studies (71 effects) • Experimental and QES • Published & unpublished • Search range: 1980-2012	• PreK-6 • Countries: 60% English-speaking • Settings: NR • Sample size: NR	• Reading interventions [f] • Control: NR • Duration: NR • Population Focus: Universal and targeted	• Literacy [g]	• Literacy: post test: 0.37 (ns) (#ES=71), follow up d=0.22 (ns) (#ES=71) Grade: Pre-K d=0.34 (ns) post test, 0.12 (ns), #ES:29follow up Grade 1-2; d=0.40 (ns) post-test, k=30, d=0.26 (ns) follow up , k=30 Follow up: 11.17 months	+ (74%)	Pr: 100% EYS: 83%	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review, TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately, [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Searches targeted American, Canadian, British and Australian databases to 'target material from the English speaking world'; [d] Instruction approaches include teacher led and peer assisted learning and technology. Instruction components include; advance organiser, attribution, control task difficulty, elaboration, large group, small group, one to one, peer mediation, task reduction etc; [e] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring; [f] Phonemic awareness, phonics, fluency, reading comprehension, mixed interventions; [g] literacy skills, reading, comprehension, spelling composite

Domain 2: Effective differentiated teaching

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance[b]	Evidence Rating
#350 Deunk et al (2018)	• Systematic review & Meta-analysis • 21 studies • QED & RCT (#NR) • Published • Search range: 1995-2012	• Age 6-12yrs • Countries: USA, Australia, UK, The Netherlands • Settings: NR • n~22950	• Differentiation practices [c] • Control: business as usual or secondary data analysis from large scale survey studies • Duration: 12wk-3 yrs • Population Focus: Universal	• Academic performance (language and math)	Overall ES: d= 0.146*, k=21 Overall: • Between class homogenous ability grouping d=-0.065, k=3 • Within class homogenous ability grouping d=-0.007, k=6 • Computerised system as differentiation tool d=0.290*, k=6 • Differentiation broader program contexts d=0.296*, k=6 • Early years: No pooled results -> some positive effects for computerised differentiation - Follow Up: NR	++(76%)	Pr:71% EYS: 28.5%	Well Supported
#74 Graham et al (2015)	• Meta-analysis • 35 studies • 19 RCT & 16 QES (some pre-post for computer feedback) • Published & unpublished • Search range NR; dates of included papers 1975-2012	• K-8 • Countries: NR • Settings: NR • n~9457	• Writing assessment (feedback from adults, peers, self, or computers) • Control: NR • Duration: NR • Population Focus: NR	• Writing quality	• Overall:, g=0.61*, p<0.01, k=27 • Adult feedback: g=0.87*, p<0.01, k=7; Peer feedback: g=0.58*, p<0.001, k=8; Self-assessment: g=0.62*, p<0.001, k=10; Computer feedback*: g=0.38*, p=0.001, k=4, * weak study designs. • Progress monitoring /curriculum-based measurement (to inform teaching):, g=0.18, p=0.06, k=5	++ (85%)	Pr: 72% EYS: 10%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance[b]	Evidence Rating
					Follow up: NR			
#454 Lou et al (2000)	• Meta-analysis • 51 studies • Experimental • Published & unpublished • Search range: NR, dates of included papers: NR	• Elementary-post secondary (grade breakdown not given) • Countries: NR • Settings: NR • Sample sizes: NR	• Class grouping (within class ability grouping) • Control: NR • Duration: NR • Population Focus: Universal	• Cognitive outcomes	Overall (small but significantly positive): • g=0.16* #ES:103 Class grouping: small-group instruction significantly larger with cooperative learning vs other small-group methods (e.g. unstructured group work ability ((Part r^2: 0.06;) • Effects more positive for lower grade than higher (Part r2 0.05, B= -0.10). Follow Up: NR	+(70%)	No breakdown provided.	Unknown
#23 Ok et al. (2016)	• Systematic review • 13 studies • Qual, Quant, single-case, mixed-method (8 QES, no RCTs) • Published (peer reviewed) • Search range: Jan 2000-Dec 2014	• Pre-K-12 • Countries: NR • Settings: NR • n=3550	• Universal Design for Learning • Control: NR • Duration range: 1 session (20-90 min) to 1 year • Population Focus: Universal	• Academic (reading, science, social studies) • Social	• Academic ES (10 studies): small to large. • Only 2 studies of social outcomes, effect ns. • Strong effects in single-case and secondary school studies only • ES varied by outcome • 1 study using shared stories individualised for students with multiple disabilities Follow Up: NR	+ (60%)	Pr: 38% EYS: 23%	Unknown
#490 Piasta & Wagner (2010)	• Meta-analysis • 63 studies • 39 RCT/CRCT,21 QED, 3 include both • Published & Unpublished • Search range NR, dates of	• PreK-6 • Countries: Canada, France, Germany, Israel, the Netherlands, New Zealand, the UK, USA • Setting: NR • n=8469	• Alphabet training instruction • Control: NR • Duration: 120-5,793 mins • Population Focus: Universal	• Letter sound knowledge, letter name fluency, letter sound fluency, letter name knowledge and letter writing	• Letter sound knowledge (g = 0.65*, p<0.001, k= 36) • Letter name fluency (g = -0.02*, p<0.001, k=21) • Letter sound fluency (g = 0.58*, p<0.001 ; k =4), • Letter name knowledge (g = 0.43*, p<0.001, k =27), • Letter writing, (g= 0.59*, p=0.012, k= 6.)	++(78%)	Pr: 100% EYS: 92%	Well Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance[b]	Evidence Rating
	included papers 1980-2007				• Grouping effects: Significant for Letter name knowledge: small groups (g=0.52, k=21) vs individual tutoring (g=0.26, k=4) vs whole class (g=0.24, k=12)). No difference for other outcomes. **Early years** (preschool/kindergarten): letter name knowledge (g=0.37 k=33); letter sound knowledge (g =0.65, k=35); letter writing (g= 0.601, k=5); letter name fluency (g=0.09, k=14) Effects of instruction content explored. Follow Up: NR			
#46 Santangelo & Graham (2016)	• Meta-analysis • 76 papers (80 experiments) • RCTs & QES • Published & Unpublished • Search range: 1931-2015	• K to 12 • Countries: NR • Setting: NR • Sample sizes: NR	• Handwriting instruction methods • Control: non-handwriting instruction • Duration: NR • Population Focus: Universal	Handwriting legibility and fluency	Effects on legibility: • Individualized instruction g=0.69*, p<0.01, k=8 Follow up: NR	++ (87%)	Pr: 93% (overall), 89% (domain) EYS: 73% (overall), 50% (domain)	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review, TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Homogenous ability grouping, computerised differentiation, and differentiation in broader reform context; [d] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring

Domain 3: Peer tutoring and collaborative learning

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
#94 Dietrichson et al. (2017)	• Meta-analysis • 101 studies (10 for co-operative learning) • RCTs (76%) & QES • Published & unpublished • Search range: 2000-2014	• K-9 • Countries: OECD & EU (95% US) • Settings: NR • n~1.07 million	• Interventions to improve academic achievement (within existing systems)[c] • Control: waitlist, placebo • Duration: mean 30 weeks • Population Focus: Universal	• Academic (standardised reading and math tests)	Overall ES (both small but significant): • Reading g=0.09*, k=66 RCTs • Mathematics, g=0.08*, k=25 RCTs Cooperative learning: g=0.22*, k=10 Follow up: No restrictions	++ (85%)	Pr: 78% (overall), NR (domain) EYS: NR	Unknown
#361 Elbaum et al (1999)	• Meta-analysis • 20 studies • Study type: NR (quantitative) • Published & unpublished • Search range: 1975-1995	• Grade 1-6 • Countries: NR • Settings: Urban, suburban, rural • n=573	• Grouping students for instruction • Control: NR • Duration: 5-36.5 hours • Population Focus: Targeted (Learning disabilities & behavioural disorders)	• Reading (decoding, fluency, comprehension • Writing • Spelling	ES= glass Δ Overall: 0.40*; #ES=70 • Decoding: 1.02*; #ES= 7 • General reading: 0.59*; #ES =7 • Reading comprehension: 0.41*; #ES =21;oral reading of words: 0.27*; #ES = 16, oral reading of passages: 0.09 (ns); #ES= 11, spelling: -0.05 (ns), #ES: 4, composition/writing: 0.35(ns); #ES=3 and language mechanics: 0.14 (ns); #ES=1 • Peer tutoring: 0.24* #ES=13 vs cross age tutoring: 0.50*; #ES=12 vs co-operative partners: 0.00(ns); #ES = 1 • Breakdown of results by grade not given Follow up: NR	+(57%)	Pr:100% EYS: NR	Supported
#430 Kunsch et al (2007)	• Meta-analysis • 17 studies • 16 experiments, 1 QED • Published	• K-12 • Countries: NR • Setting: 13/17 =	• Peer mediated instruction • Control: NR	• Maths performance	• Overall: d=0.47, k=17 • Elementary-aged students d=0.57, SD: 0.41; k=14. • ES for disabled children (d=0.21, SD 0.23, k=10) vs at	+ (74%)	Pr 82% EYS: 18%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	• Search range: 1978-2006	general education; 4= special education • n=3046	• Duration: 8-45 mins, 2-3 times/week • Population Focus: Targeted (Students with/at risk of disability)		risk children (d=0.66, SD 0.42, k=12) vs combined (d= 0.53, no SD, k=1) • Treatment components: • EYS: NR (but d=0.57 Primary, k=14, and 3 individual studies presented Follow Up: 4wks-1 year			
#443 Leung (2015)	• Meta-analysis • 72 studies • 19 QED, 52 experimental 1 mixed mode • Published • Search range: NR, dates of included studies 1990-2012	• K-Tertiary • Countries: NR • Setting: 72% high SES • Sample size: NR	• Peer tutoring • Control: NR • Duration: > 16.25 hrs =29.2%, ≤16.25hrs 30.5%, remainder=NR • Population Focus: Universal	• Academic performance (maths, reading, language, science & tech, PE, Arts, psychology, education)	Mixed effects, with trim & fill analysis, weighted d: • Overall d=0.37*, k=72 • Grade: Elementary school (d =0.47*,p<0.001; k= 46), and kindergarten (d =0.21, ns, k=2) Moderators: • Structured tutoring: larger ES (d =0.53*, k =53) vs. unstructured tutoring (d =0.33*, k=23). • Reward type: tangible items (d = 0.70*, k=12) vs. points as rewards (d =0.35*, k= 9). • Gender of the dyads: same-gender (d = 0 .80*, k=14) vs mixed dyads (d = 0.41*, k= 58). Follow Up: NR	+ (70%)	Pr: 67% tutees; 60% tutors EYS: NR	Supported
#110 Puzio (2013)	• Meta-analysis • 18 studies of 29 unique cohorts • 2 CRCT, 6 MC, 10 QES • Published & unpublished • Search range: 1980+; included	• Grade 2-12 • Counties: 17 USA, 1 other • Setting: Urban, suburban, rural	• Within-class cooperative or collaborative grouping • Control: NR • Duration: < 10 weeks to full school year • Population Focus: Universal and	• Reading • Vocabulary • Comprehension	Positive effects, hedges g: • Reading =0.16 , p=0.001 • Comprehension = 0.20, p<0.001 • Vocabulary =0.22, p<0.001 • Follow Up: NR	++ (85%)	Pr: 89% EYS: NR	Well Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	studies: 1987-2009	• Sample sizes: 12,286	targeted (Regular and Special Ed.)					
#497 Rohrbeck et al (2003)	• Meta-analysis • 90 studies • QED & Experimental (#NR) • Published • Search range NR; included studies 1990-2012	• 1-6 • Countries: NR • Setting: 26 urban, 19 suburban, 7 rural • Sample size: NR	• Peer assisted learning strategies [d] • Control: NR • Duration: 3 to 1,080 hrs. • Population Focus: Universal	• Student achievement (Reading, mathematics, social studies, science, writing, language, literacy)	Overall: d=0.33*, p≤0.0001, k=81. • Reading d= 0.26*, p≤0.0001, k= 26; math d=0.22*, p≤0.0001, k=33; social studies d=0.49*, p≤0.0001, k=13; science d=0.62*, p≤0.0001, k=6; spelling d=0.21*, p≤0.01, k=3; writing d=0.33*, p≤0.0001, k=5; language d=0.21*, p≤0.0001. k=11, and literacy, d=0.27*, p≤0.0001 k=35 • Tutee grades 1-3 d=0.37*; k=26 vs grades 4th-6th ES 0.28*, k= 44. Follow Up: NR	+(63%)	Pr: 100% EYS~46%	Supported
#577 Zenelli et al (2016)	• Meta-analysis • 41 studies • 17 PP, 3 QED, 29 RCTs • Published • Search range NR, included studies 1965-2014	• Elementary & high school (4-18 years) • Countries: 21 USA, 17 UK, 3 NR • Setting: NR • Sample size: NR	• Peer tutoring • Control: NR • Duration: 20; 7-12 weeks, 11; 6 weeks, 10; 1-52 weeks • Population Focus: Universal	• Child academic achievement (maths, reading science)	• Control -group Studies; d= 0.25*, #ES=32 • Grade Level: Control-group studies. Elementary d=0.51*, k =30 Follow Up: NR	+(69%)	Pr: 88% EYS: NR	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review, TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring; [d] Includes investigation of those using reciprocal interaction, and reward contingencies.

Domain 4: Physical activity for academic achievement

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
#352 Donnelly et al (2009)	• CRCT (24 schools) • 14 intervention schools vs 10 control schools	• Grade: 2-3 • Country: USA (Northern Kansas) • Setting: NR • n= 1527 (IG:814, CG:713) • Implementer: classroom teacher	• PAAC- Physically active academic lessons[c] • Control: Regular classroom instruction • Dose/Duration: 90mins/wk • Population Focus: Universal	• BMI • daily PA • academic achievement	Positive program effects for achievement (all ps<0.01): • Composite • Reading • Maths • Spelling scores No significant change in BMI. Positive program effects for PA (p<0.01) Follow up: NR	+ (73%)	100% EYS	Promising
#353 Donnelly et al (2017)	• CRCT (17 schools) • 9 intervention schools vs 8 control schools	• Grade 2-3 • Country: USA (Eastern Kansas) • Setting: 70% urban • n=584 (IG: 316 vs CG: 268) • Implementer: classroom teacher	• A-PAAC – Classroom teachers trained to deliver academic lessons using moderate to vigorous physical activity in a subject of their choice • Control: No intervention • Dose/Duration: ≥ 100mins/week A+PAAC lessons + physical education (60mins) = 160min • Population Focus: Universal	• Academic achievement • Anthropometry[d]/ • cardiovascular fitness	Positive program trends: • Reading (p=0.056); • Math(p=0.082) • No effect for spelling (p>0.10). No significant effect on BMI, waist circumference or cardiovascular fitness Follow up: NR	+ (56%)	100% EYS	Not Supported
#396 Harvey et al (2018)	• RCT (2 schools)	• Grade: 2-3 • Country : USA (Northeast Kansas) • Setting: Urban public school	• A+PAAC – As above • Control: Lessons delivered in usual manner • Dose/duration: 100min of PA +	• Learning/academic behaviour	Behaviour engagement over time: • significantly improved for intervention; (T1 to T2 ; d= 0.43*, from T2 to T3 d = 0.81*) • no change or slight degradation for control (T1 to T2 = d = 0.03	+ (71%)	EYS: 100%	Preliminary

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
		• n= 68 (IG: 43 vs CG: 25) • Implementer: Classroom teacher	PAAC activities per week • Population Focus: Targeted (students with learning difficulties)		(ns), from T2 to T3 = d = 0.16 (ns). Follow Up: NR			
#444 Li et al (2004)	• RCT	• Grade: 2 • Country: NR • Setting: NR • n=83 (IG: 42 vs CG: 41) • Implementer: PE teacher	• Sport stacking – stacking cups to form pyramids • Control: Supervised other physical activity • Dose/duration: 15min/school day for 14 weeks • Population Focus: Universal	• Handwriting quality and speed	• Writing speed: positive trend, d=0.32, p=0.08 (ns) • Letter formation accuracy : ns Follow up: NR	+ (56%)	EYS: 100%	Not Supported
#477 Mullender-Wijnsm (2016)	• CRCT (12 schools)	• Grade 2-3 • Country: Netherlands • Setting: NR • n=499 (IG: 249 vs CG: 250) • Implementer: teacher	• "Fit & Vaardig op School" [F&V]": moderate-vigorous physical activity (MVPA) to teach math and spelling. • Control: Regular classroom lesson • Dose/duration: 22 weeks (3x per week for 20-30 mins) • Population Focus: Universal	• Academic achievement in maths and language	Results for 2yr follow up: • Reading: d= 0.05 (ns), p=0.45, n=181. • Spelling: d = 0.45*, p <0 .001, n=180. • Math speed: d=0.51*, p <0 .001, n=181. • General maths d=0.42*, p <0.001; n=179. Follow-up: 2 years	+ (60%)		Promising
#48 Martin & Murtagh (2017)	• Systematic review • 15 studies • RCTs, CRCTs, QES, PP, delayed	• Primary students (ages 5-12 years) • Countries: USA, China, Netherlands,	• Physically active teaching methods • Control: NR • Duration range: 5 consecutive days to 3 years	• physical activity • academic (Literacy, science and maths) • health (e.g. BMI)	• Physical activity: 6 of 10 studies (3 RCTs, 2 QES) show medium-to-large effect (ES 0.24 to 2.48, p<.05 for both)	++ (81%)	Pr: 93% EYS: 13%	Well Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	treatment, exploratory, • Peer reviewed journals • Search: Jan 1990-Mar 2015	New Zealand, Australia • Settings: classroom • n=9067	• Population Focus: Universal		• Learning outcomes: All of 4 studies (2 CRCTs) show positive effects (1 CRCT of 7-9 year olds) • BMI: 3 studies (2 CRCT) show small positive ES Follow up: NR			
#34 Owen et al. (2016)	• Systematic review & meta-analysis • 38 studies • 9 RCT, 15 QES, cohort, cross-sectional • Published & unpublished • Search: Up to Dec 2014	• 5-18 years • Countries: 17 USA, 3 Australia, Others • Settings: NR • n=71,433	• Physical Activity (PA): Integrating physical activity in academic lessons[e] • Control: usual practice, no program, alternative program • Duration: NR • Population Focus: Universal	• School engagement (behavioural, emotional, cognitive)	• Small positive association for PA and overall school engagement (d=0.28*, k=29; for children d=0.27*, k=22). • Significant effect in RCT studies (d=0.40*, k=7) • No association with disengagement (d=-0.32, k=11) • Medium association for PA breaks during class, d=0.55*, k=4. Moderators assessed: intervention type, frequency, age, risk of bias, Follow up: NR	++ (88%)	Pr: 79% - 100% EYS: NR	Well Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies;[c] The PAAC intervention was delivered by teachers who have received in-service training. The PAAC intervention involved delivery of 90minutes/wk of moderate to vigorous intensity physically active academic lessons (3.0 to 6.0 METS, apx 10 min each) intermittently throughout the school day; [d] height, weight (used to calculate BMI) and waist circumference; [e] Includes analysis of specific strategies: providing physical exercise breaks in class, single bouts of exercise in 60 min prior to class

Domain 5: Technology assisted teaching and learning

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
#58 Abrami (2015)	• Meta-analysis • 11 studies • 4 CRCT, 3 RCTs and 3 QED, 1 pre-experiment • Published & unpublished • Search range: 2008-2014	• PreK-3 • Countries: Canada, Australia, Kenya, Hong Kong • Setting: urban & rural • n=2880	• ABRACADABRA: comprehensive, interactive web-based reading software. • Control: Regular literacy instruction • Duration: 20-1920hrs • Population Focus: Universal	• Overall • Phonics • Phonological awareness • Reading fluency • Reading Comprehension • Vocabulary • Listening comprehension	• Overall g=0.179*, p<0.01 #ES=73 • Phonics g= 0.189*; p<0.01 #ES=19, phonics awareness g=0.324*, p<0.001, #ES=20, reading fluency g =0.078,ns, #ES=6, reading comprehension g=0.065, ns , #ES=6, vocabulary knowledge g= 0.108, ns, #ES=15, listening comprehension g=0.184, ns, #ES=7 Follow up: NR	+ (67%)	EYS: 100%	Supported
#72 Cheung and Slavin (2012)	• Meta-analysis • 84 studies • 25 RCTs, 3 QES, MC, and matched post hoc • Published & unpublished • Search range: 1980-2010	• K-12 • Countries: NR • Settings: NR • n>60,000	• Any educational technology to improve reading (e.g. computers, multimedia, interactive whiteboards) • Control: traditional methods • Duration: At least 12 weeks • Population Focus: NR	• Reading	Overall: small positive effect of technology vs traditional methods (d'=+0.16*, p<0.001, k=84) • CAI programs that have dominated classroom use had minimal effect (d'=+0.11*, p<0.001, k=56); • Greater effects for innovative technology applications (d'=0.18*, p<0.001, k=6) and integrated interventions with extensive PD (d'=0.28*, p<0.001, k=18) Kinder d'=0.15 (p=0.28),k=8 elementary d'=0.10*,p<0.001, k=59 Follow up: NR	++ (79%)	Pr: 79% EYS: 9%	Supported
#94 Dietrichson et al. (2017)	• Meta-analysis • 101 studies (9 using Computer Assisted Instruction)	• K-9 • Countries: OECD & EU (95% US) • Settings: NR • n~1.07 million	• Interventions to improve academic achievement (within existing systems)[c] • Control: waitlist, placebo	• Academic (standardised reading and math tests) •	Overall ES (both small but significant): • Reading, g=0.09*, k=66 RCTs • Mathematics , g=.08*.k=25 RCTs Computer-assisted instruction: g=0.11,ns, k=9	++ (85%)	Pr: 78% (overall), NR (domain) EYS: NR	Unknown

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	• RCTs (76%) & QES • Published & unpublished • Search range: 2000-2014		• Duration: mean 30 weeks • Population Focus: Universal		Follow up: No restrictions			
#440 Lee (2013)	• Meta-analysis • 58 studies • Experimental and quasi-experimental studies #NR • Published • Search range: 1997-2011	• K-12 • Countries: NR • Setting: NR • Sample size=NR	• Teaching and learning with technology[d] • Control: non-technology comparison group • Duration: NR • Population Focus: Universal	• Cognitive and affective outcomes	Overall: positive cognitive effect (d=0.42, k=48). • Grades K-3: d=0.50 • Grades 4-6: d = 0.41 • Type: PCs (d=0.56), laptops (d= 0.88), networked computer (d= 0.39) , multimedia (d= 0.61) {k NR} Follow up: NR	+ (60%)	Pr: 50% EYS: 24%	Supported
#23 Ok et al. (2016)	• Systematic Review • 13 studies • Qual, Quant, single-case, or mixed-method (8 QES) • No RCTs • Peer-reviewed journals • Search range: Jan 2000-Dec 2014	• Pre K to 12 • Countries: NR • Setting: NR • n= 3550	• Universal Design for Learning • Control: NR • Duration range: 1 session (20-90 min) to 1 year • Population Focus: Universal & targeted (disability) (k=10 inclusive classrooms; 1 both inclusive and separate; 1 special education class)	• Academic • Social	Descriptive results only: • Academic ES (10 studies): small to large. • Only 2 studies of social outcomes, effect ns. • Strong effects in single-case and secondary school studies only • ES varied by outcome • 7 studies examined technology-based environments that aligned with UDL principles (3 on digital text environments and literacy; 4 on digital materials for science and social studies). Follow up: NR	+(60%)	Pr: 38% EYS: 23%	Unknown
#46 Santangelo & Graham (2016)	• Meta-analysis • 76 studies (80 experiments) • RCTs & QES	• K to 12 • Countries: NR • Setting: NR • Sample sizes: NR	• Handwriting instruction methods • Control: non-handwriting instruction	• Handwriting legibility and fluency	Effects on legibility: • Using technology as part of handwriting instruction g=0.85*, p<0.05, (k=4, 3EYS); Effects for fluency: NR	++ (87%)	Pr: 93% (overall), 100% (domain) EYS: 73% (overall)	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	• Published & Unpublished • Search range: Up to 2014 (range 1931-2015)		• Duration: NR • Population Focus: Universal		Follow up: NR		75% (domain)	
#55 Slavin et al. (2012)	• Systematic review • 17 studies • RCTs, QES, & matched controls • Published and unpublished • Search range: 1980-2011	• K-6 • Countries: England, USA, Taiwan, Kuwait • Setting: Urban & Rural • Sample sizes: NR	• Programs and practices used in elementary science • Control: alternative programs or standard methods • Duration: 4 weeks minimum • Population Focus: NR	• Academic (science specific)	• Integrating technology (video & computer) with teaching and cooperative learning showed promising results (d'=+0.37, k=5); Follow up: NR	+ (52%)	Pr: 100% EYS: 6%	Supported
#516 Sokolowski (2015)	• Meta-analysis • 24 studies • 10 RCT, 14 QED • Published • Search range: 2010-2013	• Grade 1-8 • Countries: Cyrprus, Hong Kong, Tawiwan, USA, Turkey, Israel, Greece, Kuwait, Netherlands. • Setting: NR • n=4526	• Exploratory computerised environment (ECE), digitally delivered used to formulate and mathematize patterns or solve problems. • Control: Traditional instruction methods • Duration: 1-80 wks • Population Focus: Universal	• Student achievement	Overall: g= 0.60*, k=24 Grade Level: • 1-3: g=0.61*, k=3 • 4-5: g=0.41*, k=12 • 6-8: g=0.65*; k=9 Follow up: NR	++ (80%)	Pr: 71% EYS: 12.5%	Well Supported
#528 Sung (2016	• Meta-analysis & research synthesis • 110 studies • Experimental and QED #NR • Published	• Kindergarten - graduate school • Countries: NR • Setting: NR • n=18749	• Mobile devices as learning tools in classroom/outdoor learning[e] • Control: NR • Duration: ≤4hr - >6 months	• Student learning performance[f]	Overall: g= 0.523, #ES 419 studies. Subject: • Language arts (g= 0.59, #ES=41) • Social studies (g = 0.78, #ES = 10), • Science (g= 0.58, #ES =78), • Maths (g = 0.34, #ES =41)	+ (68.5%)	Pr: 35% EYS: 0.9%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
	• Search range NR, included studies: 1993-2013		• Population Focus: Universal		Grade: kindergarten g= 0.10, #ES 2; elementary school g=0.65, #ES 97. Follow-up: NR			
#532 Tackas 2015	• Meta-analysis • 43 studies • 42 experimental, 1 QED • Published & unpublished • Search range: 1980-2014	• Preschool and/or elementary • Countries: 24 United States 3 UK, 11 Netherlands, 5 Israel • Setting: NR • n=2147	• Technology enhanced stories for literacy development[g] • Control: Traditional - Storybook reading • Duration: NR • Population Focus: Universal	• Story Comprehension • Expressive vocabulary receptive vocabulary • Code-related skills • Engagement	• Story comprehension g+ = 0.17*, #ES=38, • expressive vocabulary g+ = 0.20*, #ES=18, • receptive vocabulary g+ = -0.08, ns, #ES=9, • code related literacy skills g+=0.16, ns, #ES=14, • engagement • child initiated communication during reading g+ =0.26, ns, #ES=12 Follow up: NR	++ (81.4%)	Pr: 100% EYS: NR	Well Supported
#76 Thomas et al. (2013)	• Meta-analysis • 40 studies • RCT &QES • Published & unpublished • Search range:1990-2011	• K-Tertiary • Countries: NR • Settings: NR • n=4925	• Interactive computer-Assisted Instruction (CAI) • Control: CAI in didactic environment • Duration: NR • Population Focus: Universal (some studies separate by ability)	• Academic (grades, subjects NR)	• g= 0.175, p<0.05, #ES 55 • Elementary level studies: g = 0.186, ns, #ES 13 • ES smaller in secondary levels, higher in tertiary. Follow up: NR	+ (67%)	Pr: 24% EYS: NR	Supported
#149 Tingir et al.	• Meta-analysis • 14 studies • 3 Experimental, 9 QES • Published • Search range: 2010-2014	• K-12 • Countries: NR (but languages were English, Chinese, Spanish & Taiwanese) • Settings: NR • Sample size=NR	• Mobile device interventions provided as regular part of curriculum • Control: Traditional teaching • Duration: NR • Population Focus: NR	Achievement (science, math, reading)	Overall: positive effect, d=0.48*, #ES=27 Elementary students d=0.55*, #ES= NR Moderators assessed: Device type, subject area, study design, implementer Follow up: NR	++ (82.4%)	Pr: 57% EYS: NR	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring; [d] E.g. use of multimedia, PDs, integrated learning systems; [e] PDAs, smart phones, laptops; [f] Language arts, social studies, science and maths; [g] E.g. multimedia and interactive features

Domain 6: Physical environment design to optimise learning

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
#305 Amlani & Russo (2016)	• Repeated Measures Design	• Grade 3 • Country: USA (Texas) • Setting: NR • n=27 • Implementer: NR	• Use of acoustic panels • Control: repeated measures-no panels • Dose/Duration: NR • Population Focus: Universal	• Speech transmission and listening effort[c]	• Word Recognition performance: 88% (95% CI ±1.3) without acoustic panel vs 79% (95% CI±1.4) with acoustic panels ($p<0.001$). • Digital recall performance: 79% (95CI±1.1) without acoustic panels vs 69% (95% CI ± 1.1) with acoustic panels ($p<0.001$). •	+(73%)	EYS: 100%	Unknown[1]
#368 Fisher et al (2014)	• Repeated Measures Design • students split into two groups to prevent overcrowding – groups matched on age and gender	• Kindergarten • Country: USA (Midwestern city) • Setting: NR • n=24 • Implementer: Female researcher with prior experience with early childhood education	• Decorated vs sparse classroom • Control: repeated measure • Dose/duration: x5 familiarisation sessions plus 6 experimental sessions • Population Focus: Universal	• Time spent off task and type of learning	• % time spent off task greater in decorated vs sparse classroom ($d = 0.85$, $p<0.001$). • % time engaged in environmental distractions lower in sparse classroom vs decorated ($d = -2.60$, $p<0.001$). • Children's learning scores higher in sparse-classroom vs decorated-classroom ($d = 0.65$, $p=0.007$).	+ (63%)	EYS: 100%	Unknown (design flaws)
#3 Gunter & Shao (2016)	• Meta-analysis • 18 studies • Quantitative designs • Published & unpublished • Search range: NR	• K-12 • Countries: NR • Settings: NR • Sample size: NR	• Condition or features (cosmetic or structural) of school buildings • Control: NR • Duration: NA • Population Focus: NR	• Academic (standardised tests)	• Overall: weighted mean correlation r(bivariate)= 0.12*, #ES 594, $p<0.001$; regression studies r(semi-partial)= 0.10*, $p<0.001$ • Elementary school: correlation studies r(bivariate) = 0.20*,$p<0.0001$ #ES 61; regression studies, r(semi-partial)= 0.11*,$p<0.0001$ #ES=11 • Correlation for subjects $p<0.05$ for all: science r=0.17*,#ES=92	+(52%)	Pr: 10% (correlation studies) 83% (regression studies) EYS: NR	Supported (RQ1)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
					language arts r=0.14*, #ES=193, math r=0.14*, #ES=92			
#487 Pfeiffer et al (2008)	• RCT	• Grade 2 • Country: USA (Pennsylvania) • Setting: NR • n=64 (IG: 32 vs CG: 32) • Implementer: Teacher	• Disc O'Sit Cushion: round air-filled cushion designed to provide movement while seated • Control: regular classroom chair • Dose/duration: 2hr/day for 2 weeks • Population Focus: Targeted	• Attention to task	• Global Executive Composite (GEC) η^2=0.324, p<0.001 • Behavioural regulation index (BRI) η^2=0.229, p<0.001. • Metacognition index (MI) scores was found η^2=0.145, p<0.01.	++(75%)	100% EYS	Promising

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] The evidence suggesting possible negative effects for acoustic panels was rated unknown, because there were serious limitations to external validity

Domain 7: Class size and teacher-student ratios

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes [a]	Main Findings	Quality Rating	Overall Relevance [b]	Evidence Rating
#100 Glass & Smith (1979)	• Meta-analysis • 77 studies • RCT, matched, repeated measures, or uncontrolled studies (47% RCT or matched) • Published & unpublished • Search range: 1900-1978	• Grade/Age: NR • >12 countries: NR • Settings: NR • n~900,000	• Class size • Control: NR • Duration: 1-9000 hours • Population Focus: NR	• Academic (standardised tests)	Overall, results favoured smaller class size: Glass delta=0.09, #ES 725 • Class-size of 1 vs class-size of 40: ES=0.57 (i.e. ~70th percentile vs ~50th percentile) • Class-size 20 vs 40: ES=0.051 overall, ES=0.063 for primary students (~10 percentile point difference) Follow up: NR	-(48%)	Pr: 56% EYS: 15%	Unknown
#112 Shin & Chung (2009)	• Meta-analysis • 17 studies • 13 RCT, 4 QES • Published & unpublished • Search range: 1989-2008	• K-12 • Country: USA • Settings: NR • Sample size: NR	• Treatment groups in small class • Control: regular or large class >22 • Duration: NR • Population Focus: NR	Academic (social science, science, math, reading, writing - standardised tests)	Overall: positive effect favouring smaller class size (vs large class size) d=0.20*, p<0.05, #ES=120 • For elementary schools d=0.20*,p<0.05, #ES:114 • secondary school d =-0.05 , p=0.8(ns), #ES:6 • Moderators: grade level, publication status, study quality • Similar ES for social science, math, reading (d=0.19 to 0.20) • Follow up: NR	+ (59%)	Pr: 94% EYS: 76%	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d, d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES=Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero; ~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring.

Domain 9: Social-emotional and behavioural (SEB) interventions to promote a positive school climate

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
#109 Barbero (2012)	• Meta-synthesis • 32 studies • MAs, SRs, RCTs, NRCTs, PP, cohorts • Peer reviewed journal publications • Search range: Jan 2000-2011	• 5-16 years • Countries: Belgium, USA, Italy, Holland, Australia, Canada, Greece, China, Germany, UK • Settings: NR • Sample size: NR	• Violence reduction programs • Control: NR • Duration: 3 days 5 years; • Population Focus: Universal	• Behavioural • knowledge and attitudes about violence and bullying at school; • academic • social-emotional • perceived safety • parental involvement	No meta-analysis. Most efficient interventions: • Have multi-disciplinary involvement • Global focus • Aim to improve social and interpersonal skills and modify attitudes and beliefs • MAs and SRs (k=5) show positive effects on bullying and victimization, social competence, self-esteem and peer acceptance • RCTs: significant effects for 9 of 12 studies (and 6 of 7 primary level RCTs). Follow up: Up to 3 years for RCTs, 5 years non-RCT	+ (73%)	Pr: 53% 58% (RCTs), 40% (MAs & SRs; EYS: 3% Primary-specific results presented.	Supported
#328 Carsley et al (2018)	• Meta-analysis • 24 studies • Quantitative (no other detail) #NR • Published • Search range: NR (included studies 2005-2016)	• Elementary-high school (6-18 yrs) • Countries: NR • Settings: NR • Sample size: NR	• School based mindfulness interventions for mental health and well being • Control: NR • Duration: 0.25-36hrs • Population Focus: Universal	• Mental health & well being	Between-group studies: • post-test (g= 0.24*, p<0.001, #ES=21), • follow-up (g=0.17, p=0.079, #ES=6). Middle childhood (6-10 years): between-group studies post-test g=0.22*, p<0.05, #ES=7. For within group at post-test g = 0.20*, p<0.05, #ES=6. Follow up: 12-32 weeks	++(89%)	Pr: 54% EYS: 4%	Supported
#94 Dietrichson et al. (2017)	• Meta-analysis • 101 studies (7 psychologic	• K-9 • Countries: OECD & EU (95% US) • Settings: NR	• Interventions to improve academic achievement (within existing systems)[d]	• Academic (standardised reading and math tests)	Overall ES (both small but significant): • Reading g=0.09*, k=66 • Mathematics g=0.08*, k=25	++ (85%)	Pr: 78% (overall), (NR domain) EYS: NR	Unknown

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	al/ behavioural) • RCTs (76%) & QES • Published & unpublished • Search range: 2000-2014	• n~1.07 million	• Control: waitlist, placebo • Duration: mean 30 weeks • Population Focus: Universal		Psychological/behavioural interventions: g=0.05, ns, k=7 • Follow up: No restrictions			
#959 Durlak et al 2011	• Meta-analysis • 213 studies • RCTs (47%) & QES (53%) • Published & unpublished • Search range 1970-2007	• K-high school • Countries: 87% USA, 13% other • Setting: 47% urban, 16% suburban, 15% rural, 14% combination, 8% NR • n= 270,034	• Social and emotional learning programs • Control: NR • Duration: Mean no. of sessions= 40.8, 77% programs <1 yr, 11% 1-2 yrs, 12% >2 years • Population focus: Universal	• Emotional skills, attitudes, behaviour, academic performance	Post-test: Overall g=0.30*, k=213 SEL skills g=0.57*, k=68 , Attitudes g=0.23*,k=106, Prosocial behaviour g=0.24* k=86 Conduct problems 0.22*, k=112 Emotional distress g=0.24*, k=49, Academic g=0.27*, k=35 Follow up (median 52 weeks): SEL skills g =0.26*, k=8 Attitudes g =0.11*, k=16 Prosocial behaviour 0.17* k=12 Conduct problems 0.14*, k=21 Emotional distress 0.15*, k=11 Academic g = 0.32*, k=8	+(65%)	Pr: 56% EYS: NR	Supported
#19 Korpershoek et al. (2016)	• Meta-analysis • 47 studies (54 classroom management interventions) • Experimental & QES with control	• Pre K-6 • Countries: 72% USA, remainder Canada and EU • Settings: NR • Sample size: NR	• Classroom management strategies or programs • Control: no treatment or usual practice • Duration: 6 weeks to 3 years • Population Focus: Universal	• Academic • Behavioural • Emotional • Social • Motivational	• Small significant effect overall (g =0.22, p<0.01) • Small significant effects: academic, social-emotional, and behavioural (g= 0.17, p<0.01; g=0.21, p<0.01; g= 0.24, p<0.01) • Motivational (g=0.08, ns) • Pre-K to Grade 1: significant overall (g=0.28, p<0.01) and on academic, social-emotional and behavioural outcomes (g=	++ (78%)	Pr: 78% EYS: 41%	Well Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	• Published • Search range: 2003-2013				0.23, p<0.01; g= 0.25, p<0.01; g=0.27, p<0.01) • Grade 2 to 6 results similar • Key components examined • Follow up: 2 Studies only (6 weeks to 3 years) No k/#ES given.			
#166 Maynard, et al. (2012)	• Meta-analysis • 28 studies • 16 RCT/QED, 12 single group pre-post • Published & unpublished • Search range: 1990-2009	• Elementary & secondary (RCT/QED Mean =13.73 years) • Countries: All RCT/QES: USA • Others: UK, Australia, Canada • Settings: NR • n=2453; 1725 RCT/QES (CG:902 vs IG:823)	• Truancy programs (75% school-based for RCT/QES) • Control: treatment as usual, no treatment, waitlist, alternative • Duration: 1-72 weeks (RCT/QES studies) • Population Focus: Targeted (chronic absenteeism)	• Attendance Secondary outcomes: • Academic • Attitude to school • Social-Emotional • Family functioning	• Moderate positive attendance effect in RCT/QED studies (g =0.46*, p<0.05, k=16) • For elementary students RCT/QED studies attendance g= 0.16, ns, k=2 • Data for secondary outcomes too limited for analysis	++ (85%)	Pr: 25% EYS: NR	Unknown/ Not Supported (primary)
#81 Sklad al. (2012)	• Meta-analysis • 75 studies • 42 Experimental & 33 QES with controls • Published & unpublished • Search range: 1995-2008	• Primary & secondary (mean age=10.5 years; SD=2.1) • Countries: North America (~75%), Europe (15%),others NR • Settings: NR • Sample size: NR	• School-based social, emotional, or behaviour programs • Control: NR • Duration: 1-day workshop to 6 years • Population Focus: Universal	• Social • Emotional • Behavioural • Substance use • Mental health • Academic	• Small-moderate positive post-test effects on all seven outcomes (d= 0.09 -0.70) • small significant benefits on all seven outcomes at follow-up (d=0.07 -0.26) • Positive effects for primary students on increased social skills (d=0.67, p<0.001) and decreased antisocial behaviour (d=-0.59, p<0.001) • Follow up: post-test (0-6 months); follow up (7-18 months); follow up (19+ months)	+ (59%)	Pr: 41% EYS: NR Results for primary presented separately.	Supported
#618 Ttofi et al	• Systematic review and	• K-high school	• Intervention vs Control	• Bullying and victimisation	Overall Odds Ratio: • bullying 1.36*, k = 38,	+(57.4%)	Pr: 43% EYS: NR	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	meta-analysis • 44 program evaluations • 14 RCT, 21 QED , 9 age cohort design • Published & unpublished • Search range: 1983-2009	• Countries: 5 USA/Canada vs. 29 other • Settings: NR • Sample size: NR	• Anti-bullying program • Control: No anti-bullying program • Duration: ≤40 days (23 studies),≥270 days (20 studiess) • Population Focus: NR		• victimisation 1.29*, k=38. Age/Grade (≤10 years): • bullying; OR= 1.22 (k=18), • victimisation OR 1.22 (k=18)			
#51 Whear et al. (2013)	• Systematic review • 14 studies • 8 RCTs, 6 QES • Published & unpublished • Search range: Up to Sep 2011	• 2.75 to 12 years • Countries: 8 USA, 5 Europe, 1 UK • Settings: NR • n=4614 children; 424 teachers	• Teacher training in child socio-emotional skills and/or behaviour • Control: NR • Duration: NR • Population Focus: Universal	• Social • Behavioural • Academic • School engagement • child/teacher relationships	• Statistically significant positive effects for 20 outcomes. • Significant effects for 8 of 12 studies with behavioural outcomes (7 in a positive direction). • Few effects for social, emotional and academic outcomes • ES range: g = -0.17* to 1.88* Follow up: NR	++ (92%)	Pr: 86% EYS: 64%	Well Supported
#578 Zenner et al (2014)	• Systematic review & meta-analysis • 24 studies (10 RCT, 8 QED, 5 non controlled, 1 two-arm study) • Published & unpublished • Search range: NR	• Grade 1-12 • Countries: 14 Nth.America, 7 Europe, 1 Australia, 2 Asia • Settings: Urban, suburban, rural • n= 2224	• Mindfulness interventions • Control: NR • Duration: 160-3700 mins, median of 420 mins • Population Focus: Targeted (low performing/"at risk")	• Psychosocial • cognitive • emotional problems, stress & coping, resilience[e]	Between groups: Overall: g=0.40*, p<0.001,#ES=19; Cognitive performance g = 0.80*, #ES=7; stress g = 0.39*, #ES=7; resilience g = 0.36*, #ES=13 (p for all <.05); emotional problems g = 0 .19, ns, k=9; Positive results for EYS in 2 RCTs 0.20 to 0.48, (ESs not pooled)	++(87%)	Pr: 38% EYS: 8.3%	Promising

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	(included studies 2005-2012)							
#582 Zoogman et al (2015)	• Meta-analysis • 20 studies • 13 RCT, 1 QED, 6 PP • Published • Search range: NR (included studies 2004-2011)	• Grade not provided (<18yrs) • Countries: NR • Settings: NR • n= 1914	• Mindfulness Intervention • Control: Active/wait list control • Duration: 2-24 weeks • Population Focus: 4 clinical & 16 non-clinical (16)	• Psychological & non-psychological symptoms, attention & mindfulness measures (objective & non-objective measures)[f]	Overall del[g]=0.23*, p<0.0001, k= 20 • Objective measures del=0.23*, p= 0.0006, k=6 • Non-objective measures del=0.25*, p<0.0001, k=8 • Psychological symptoms del=0.37*; p<0.0001, k=15 • Not psychological symptoms del=0.21*; p<0.0001, k=15 • Attention and mindfulness measures del=0.28*; p= 0.009 , k=6 Age (Continuous): ns	++(80%)	NR	Unknown

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d; d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES= Effect size; EYS=Early Years of School; #ES=number of ES; g=hedges g; IG=Intervention group; k=number of studies, K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; *95% CI **does not** encompass zero; ~=approximately; [a]Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies;[c] The authors report TauU 0.82 is approximately equivalent to Cohen's d = 1.99; [d] Academic interventions – such as incentives, after school programs, summer programs, coaching students, psychological interventions, personnel development, increased resources, computer assisted instruction, content changes, coaching personnel, cooperative learning, small group instruction, feedback and progress monitoring; [e] Cognitive performance, emotional problems, stress and coping, resilience, "third person ratings" (In the domain of third person ratings, parent and teacher questionnaires were grouped, dealing with aspects such as aggressive or oppositional behaviour, social skills, emotional competence, well-being, attention, and self-regulation); [f-] Objective measures (psychophysiological measures, attention and behavioural tasks), non-objective measures (teacher-,parent-,or child-report); [g] Del is a measure of the difference in pre-post ES between groups, in this case comparison between mindfulness interventions and alternative treatments

Domain 10: Teacher-student relationships

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance [b]	Evidence Rating
10a: Included systematic reviews and meta-analyses								
#342 Cornelius-White (2007)	• Meta-analysis • 119 studies • Controlled vs uncontrolled: #NR • Published • Search range: NR (included studies 1948-2004)	• PreK-20 • Countries: USA, Philippines, Brazil, Germany, UK, Canada • Settings: Urban & Suburban • n=355,325	• Classical person-centred education • Control: NR • Duration: NR • Population Focus: Universal	• Student outcomes	• The correlation between positive teacher-student relationships and positive student outcomes r=0.36* • No breakdown by grade Follow up: NR No k/n size given.	+ (74%)	NR (majority students grade 1-12)	Unknown
#19 Korpershoek et al (2016)	• Meta-analysis • 46 studies • QED and experimental #NR • Published • Search range: 2003-2013	• Grade 1-6 • Countries: 72% US, 28% Europe & Canada. • Settings: NR • Sample size: NR	• Classroom management strategies and programs. • Control: NR • Duration: 11.1% <13 wks, 55.6% 13wks-1 yr., 33.3% >1 yr. Range: 6 wks - 3 years. • Population Focus: Universal	• Academic, behavioural and socioemotional outcomes	• k= 2 interventions: teacher-student relationship. Overall g=0.13 , ns • Academic g=0.24, p<0.01, • Behaviour g= 0.06, ns, • Socio-emotion g= 0.06, ns, • Motivational g= 0.08, ns, Grade: • Pre K & 1; g= 0.28; p < 0.01 • Grade 2-6; g=0.17;p < 0.01 Follow up: NR No k/n size given.	++ (80%)	Pr: 100% EYS: 41%	Well supported
#441 Lei et al (2016)	• Meta-analysis • 57 studies • Correlational	• Kinder-high school • Countries: 4 Eastern & 53	• Examined links between affective teacher—student relationships (TSRs) and	• Students externalising behavioural problems (EBP)	Significant -ve correlations between +ve indicators of affective TSRs and EBPs (r = −0.26*, p <0 .001, #ES = 78.	+(69%)	Pr (3-12yrs): 85% EYS (3-6-yrs): 32%	Supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	• NR if published/unpublished • Search range: 2000-2016	Western cultures • Settings: NR • n=79,333	students' externalizing behaviour problems (EBPs) • Control: Nil • Duration: NR • Population Focus: Universal.		• Age 6-9: r= -0.28*, #ES=17 Age 9-12:r=-0.23*, #ES=23 • Kinder: r = - 0.19*, #ES=25 Significant +ve correlations between -ve indicators of affective TSRs and EBPs (r = 0.55*,p < 0.001,#ES = 71. • Age 6-9: r=0.56*, #ES=20 • Age 9-12: r=0.69*,#ES=14 • Kinder: r=0.48*, #ES=31. Follow up : NR			
#498 Roorda et al (2017)	• Meta-analysis • 189 studies (Correlational) • Published • Search range: 1990-2016	• Preschool-high school (12th grade) • Countries: 111 US, 78 US • Settings: NR • n=249,198	• student engagement as mediator for association between affective teacher–student relationship and student achievement. • Control: NA • Duration: NR • Population Focus: NR	• Measured student achievement	0.14* +ve r/ships -0.12* -ve r/ships (total effects)[c] Primary school studies (k=105): Direct effect of positive teacher student relationship on achievement β=0.07*; Direct effect of negative teacher student relationship on achievement β=-0.07* Follow Up: NR	+(52%)	Pr: 55.6% EYS: NR	Supported
##552 Vandenbroucke et al (2018)	• Meta-analysis • 28 studies (ES calculated for 23) • Published & unpublished • Date search range: 2009-2017	• Age: 2-12yo – mean 5.24 • Countries: Belgium, Chile, Ecuador, Italy, the Netherlands, Portugal, Turkey, USA • Settings: NR • n= 19,906	• Association between teacher - student interactions for children's executive functions[d]. • Control: Nil • Duration: NR • Population Focus: Universal	• Children executive functions (working memory, and inhibition, cognitive flexibility)	• Overall r=0.09*, p<0.001; k=23 • General executive functioning r=0.11*, p<0.001 k=3 (fixed) • Working memory: r=0.09*, p<0.01; k=7 • Inhibition: r=0.08*, p<0.05, k=17 • Cognitive flexibility: r =0.00 (ns), k=3	++(85%)	Pr:100% EYS: 96%	Well supported

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
					• Between study effect of age: β= 0.02, p=0.002 larger r in samples with higher mean age Within study effect of age was β= 0.13, , p=0.002 Follow up: NR			
10b: Included studies from search for Randomised Controlled Trials								
#596 Abry et al (2013)	• CRCT 24 schools (IG: 13 vs CG: 11) • Teachers: n=239 (IG: 132, CG:107) •	• Grade: 3-4 • Country : USA • Settings: NR • n= 24 schools, students: NR	• Responsive classroom (RC) approach • Control: Business as usual • Dose/duration: 2 RC training sessions, 3 in person consultations with RC coaches. • Population Focus: Universal	• Teacher-student interaction quality (Obs)	• Fidelity of Implementation (FOI) predicted improvements in post-test teacher-student interaction quality, β = 0.52, p=0.001. • Indirect effect of RC on Interaction Quality via FOI, β =0.85, p=0.002 • A significant negative direct effect for RC training on post-test teacher student interaction quality, β = -0.68, p=0.03 Follow up: NR	+(71%)	EYS: 100%	Unknown
#598 Capella et al (2012)	• RCT • Schools: 5 • Teachers: n= 36 (IG: 18 classes, CG: 18 classes)	• K-5 • Country: NR • Settings: Urban • Students: n= 347 (IG: 169, CG: 178)	• BRIDGE: teacher consultation & coaching program. • Control: Reflective teaching • Dose/duration: ~4.5 observation & coaching sessions, and 3.5 consultation meetings (25-30 minutes each).	• Teacher-student relationship (Student-Teacher Relationship Scale) (TR) • Others [g]	Teacher-student relationship closeness (b = 2.75, p <0 .05, ES =0 .47). Time of post-test : end-academic year.	++(79%)	Pr: 100% EYS: NR	Promising

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings [a]	Quality Rating	Overall Relevance[b]	Evidence Rating
			• Population Focus: Targeted (behaviourally challenged)					
#599 Fernandez et al (2015)	• RCT • Teachers: n=12	• K-1 • Country: Manhattan, USA • Settings: Urban public school • Students: n=118 (IG: 64, CG: 54)	• Teacher Child Interaction Training (TCIT) • Control: Non-TCIT • Dose/Duration: ~ 11 weeks training /15 sessions • Population Focus: Universal	• Teacher-rated Student behaviour • Teacher distress (Obs) • Teacher satisfaction and • Teacher skill acquisition (Obs)[h]	• Teacher skills (Obs); post test 0.48-1.34, follow up 1.08-1.50 • Teacher distress (Obs); post test 0.90*, 0.62* • Student behaviour (TR) post test 0.30 (ns), follow up 0.22 (ns) • Follow up: 1 month after TCIT ended)	+(63%)	EYS: 100%	Promising
#600 Spilt et al (2012)	• RCT • Teachers: n= 32 • Students: n= 64	• K • Country: Holland • Setting: NR	• Relationship-focused reflection program (RFRP) • Control: Interpersonal skills training • Dose/duration: two blocks of two individual sessions with a consultant, • Population Focus: Targeted (Behaviourally at risk children)	• Child behaviour, teacher-child relationship, observed teacher behaviour, teacher efficacy	No significant results for RFRP-slope Teacher reported closeness: (beta = 0 .034, ns) Teacher reported conflict: (beta = 0 .024, ns) Observed teacher sensitivity: (beta = .093,p < 0 .10), Observed behavioural management quality (beta = 0.034, ns)	+(73%)	EYS: 100%	Not Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d; d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES= Effect size; #ES=number of ES, EYS=Early Years of School; g=Hedges g, IG=Intervention group; k=number of studies, K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; Obs=Observed; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual=Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review, WWC=What Works Clearinghouse; TR=Teacher-rated; Tx=Treatment; * 95% CI **does not** encompass zero ;~=approximately; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Effects for Total model (i.e. summing direct relationship from relationship to achievement and indirect effect from relationship to achievement via student engagement);[d] Dimensions of teacher-student interactions include closeness, conflict, classroom organization; [e] No information on treatment parts targeting teacher student relationship

or interaction - rather it states how teacher-student relationship was measured (Student-Teacher Relationship Scale-Short Form); [f] Student aggression/conduct, self-regulation, behaviour, emotions [g] Academic self-concept, and peer victimization;[h] Labelled praise, behaviour descriptions, reflections

Domain 11: Staff and leadership development

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Ranking
#90 Blank & Alas (2009)	• Meta-analysis • 16 studies • 6 RCT, 10 QES • Published & unpublished • Search range: Jan 1986- Aug 2007	• K-12 (science & math teachers) • Country: USA • Settings: NR • n= 17053	• In-service teacher PD interventions (content focus) • Control: NR • Duration: 2-542 hours • Population Focus: NR	Academic	• Significant positive effect on mathematics (d=0.13* to 0.21*) • No effect on science (d=0.05 to 0.18, ns). • For primary grades, d=.27* to .32* • Study quality moderated results. Follow up: 1 day to 16 months	++ (80%)	Pr: 43% EYS: 6% Results for primary separated.	Well Supported (RQ2)
#94 Dietrichson et al. (2017)	• Meta-analysis • 101 studies (Coaching/mentoring =10, Personnel development =8) • RCTs (76%) & QES • Published & unpublished • Search range: 2000-2014	• K-9 • Countries: OECD & EU (95% US) • Low SES • Settings: NR • n=~1.07 million	• Interventions to improve academic achievement (within existing systems) • Control: waitlist, placebo • Duration: mean 30 weeks • Population Focus: Universal	• Academic performance (standardised reading and math tests)	Overall ES (both small but significant): • Reading, g=0.09*, k=66 RCTs • Mathematics g=.08*, k=25 RCTs Personnel Development • g= 0.07,ns, k=8, Coaching/mentoring personnel • g= 0.16*, k=10 Follow up: no restrictions	++ (85%)	Pr: 78% (overall), (NR domain) EYS: NR	Unknown
#153 Dunst, et al (2015)	• Meta-synthesis • 15 studies • SR, MA, other reviews of RCTs, QESs, & pre-post studies • Published & unpublished • Search range NR; dates of	• Pre K-12 • Countries: NR • Settings: NR • n=50,000 teachers & students	• In-service PD[c] • Control: NR • Duration: NR • Population Focus: NR	• Academic • Behaviour • Teacher outcomes[d]: (attitudes & beliefs, content knowledge, instructional or behavioural practices)	• 13 syntheses explicitly identified PD components for positive outcomes • Core components identified[e] • 11 syntheses on instructional or behavioural PD • 7 syntheses with most training in teachers'	+ (73%)	Pr: 33% to 73% (5/15 early childhood to grade 6; 11/15 in K-12 range) EYS: 20% (3/15)	Supported (RQ2)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Ranking
	included studies 1987-2014				classrooms, schools or pre-school settings Follow up: NR			
#603 Jacob et al (2015)	• CRCT • 126 schools (63 schools IG)	• Grade: 3-5 • Country: USA • Setting: Rural • n=126 principals, 1546 teachers, ~300 students per school	• Intervention: Balanced Leadership Program (BLPD) • Control: Business as usual • Duration: Ten x 2-day PD sessions • Population Focus: Universal	• Principle leadership, instructional climate, principal efficacy, staff turnover, student achievement.	No impact on student achievement: • Maths : Grade 3 ES=0.04, ns • Reading: Grade 3 ES=- 0.02,ns Sig effects for Principal self-ratings on: • Efficacy (0.55*), school climate (0.34*), collective differentiated instruction (0.53*) but not leadership or collaboration All teacher-ratings of Principal not significant	++ (75%)	NR	Not Supported (RQ2)
#427 Kraft et al (2018)	• Meta-analysis • 60 studies • 56 RCT, 4 QED • Published & unpublished • Search range: 2006-2017	• Grade: PreK-high school • Country: 55 US, 2 Chile, 3 Canada • Setting: NR • Sample size: NR	• Intervention: Effect of teacher coaching on student achievement [f] • Control: Business as usual • Duration (PD): ≤20 - ≥60hrs • Population Focus: Universal	• Student achievement (standardised achievement tests)	• Effect of teacher coaching on student achievement: 0.07 ns, k=16 to 0.20*, k=26 • Teacher instruction: 0.49*, k=43 • PreK: 0.11*, k=10	+(69%)	Pr: 85% EYS: NR	Supported (RQ2)
#73 Leithwood & Sun . (2012)[g]	• Meta-analysis • 79 studies • Quantitative (correlation) • Unpublished • Search range: NR	• K-12 • Countries: NR • Settings: NR • Sample sizes: NR	• Transformational School Leadership (TSL) • Control: NR • Duration: n/a	• Academic performance (typically state-wide tests)	• Association only • Small significant positive relationships: overall weighted mean r = .09*; reading (.15*), math (.18*)	-(48%)	NR	Unknown (RQ1 only)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Ranking
			• Population Focus: Universal and targeted		Follow up: NR			
#55 Slavin et al. (2012)	• Systematic review • 17 studies • RCTs, QES, & matched controls • Published and unpublished • Search range: 1980-2011	• K-6 • Countries: England, USA, Taiwan, Kuwait • Settings: Urban & Rural • Sample sizes: NR	• Programs and practices used in elementary science • Control: alternative programs or standard methods • Duration: 4 weeks minimum • Population Focus: NR	• Academic performance (science specific)	• Inquiry-based programs emphasizing PD showed positive outcomes (d'=+0.30, k=8) Follow up: NR	+ (52%)	Pr: 100% EYS: 6%	Supported (RQ2)
#51 Whear et al. (2013)	• Systematic review • 14 studies • 8 RCTs, 6 QES • Published & unpublished • Search range: Up to Sep 2011	• 2 years to 12 years • Countries: 8 USA; 5 Europe; 1 UK • Settings: NR • n=4614 children; 424 teachers	• Teacher training in child socio-emotional skills and/or behaviour through ordinary school experience • Control: NR • Duration: NR • Population Focus: Universal	• Social • Behavioural • Academic • School engagement & child/teacher relationships	• 21 outcomes had significant effects in desired direction. • ES range: g = -0.17* to 1.88* • Overall, some improvement in some outcomes, but no significant results for academic outcomes. Follow up: NR	++ (92%)	Pr: 86% EYS: 64%	Well Supported (RQ2)

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d; d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES= Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns=not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; * 95% CI **does not** encompass zero ; ~=approximately, [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Eleven research syntheses included studies of in-service professional development to promote use of different types of instructional or behavioural practices, two research syntheses included studies to promote teacher understanding and use of content knowledge or skills, and two research syntheses included studies of in-service training to promote teacher or practitioner use of different job-related practices or to support teachers' confidence in their teaching practices. The content areas of in-service training included mathematics or science (n = 5 reviews), teacher-child interactions (n = 1 review), teacher praise (n = 1 review), teacher confidence (n= 1 review), or a mixture of different content knowledge and practice (n= 7 reviews); [d] Teacher attitudes & beliefs, content knowledge, instructional or behavioural practices; [e] Sufficient duration and intensity, extended follow-up supports and opportunities to reinforce content knowledge or practice; [f] Types: examined content-specific, reading-specific, and general coaching; [g] Likely about principals

Domain 12: Partnerships with Families

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
#366 Erion (2006)	• Meta-analysis • 37 studies • 17 single subject design, 20 group design (experimental) • Published (13), unpublished (24) • Search range: 1970-2004	• Grade K-6 • Countries: NR • Settings: NR • Sample size: 1408 group design (IG: 781 vs CG: 627)	• Parental involvement • Control: NR • Duration: NR • Population Focus: Universal	• Academic (reading, spelling, math, written expression)	• Overall: g=0.55* , k=32 • Use of written instruction, modelling, supervised practice, consultation and monitoring did not significantly affect outcomes. Longer training was significantly related to outcomes g=0.67*, k=4 • Grade: K-3: g= 0.57*, k=21,4-6: g=0.48*, k=4	++(80%)	Pr: 100% EYS:88%	Supported (RQ2)
#75 Jeynes (2012)	• Meta-analysis • 51 studies • Experimental, cross-sectional, correlational designs • Published & unpublished • Search range: NR, dates of included studies 1964-2006	• Pre K-12 • Countries: NR • Settings: Urban • n=13,000	• Association • Parental involvement programs[c] • Control: NR • Duration: NR • Population Focus: NR	• Academic	Correlational overall (but results presented as d) • Parental involvement significantly associated with achievement (for pre-elementary d=0.30*, p<0.01, ; and elementary school d=0.29*, p<0.01,) • Association of involvement and achievement significant in four program types: shared reading (d=0.51*,p<0.01) emphasized partnership (d=0.35*, p<0.01), checking homework (d=0.27*,p<0.05), parent-teacher communication (d=0.28*, p<0.05), head start program (d=0.22,ns) , ESL teaching (d=0.22,ns) K size/#ES: NR • Follow up: NR	++ (80%)	Pr: 73% EYS: 47%	Supported (RQ 1 only)
#53 Kim and Hill (2015)	• Meta-analysis • 52 studies	• K-12 • Countries: NR	• Association • specific strategies parents use to	• Academic	Correlation: • Parental involvement and achievement positively related, for both fathers (r=0.14*,	++ (85%)	Pr: 55% EYS: NR	Unknown (RQ1 only)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
	• Naturalistic or experimental, longitudinal studies, cross-sectional • Published & unpublished • Search range: 1980-2013	• Setting: Urban, suburban, rural • n=52,085 father-child dyads; 65,534 mother-child dyads	enhance children's academic outcomes • Control: NR • Duration: NR • Population Focus: Universal (Excluded very special needs)		p<0.001) and mothers (r=0.15*, p<0.001) • For K-6 children, r=0.10*, p<0.05 for father studies (k=12) and r=0.07* p<0.05 in mother studies (k=23) • Follow up: NR			
#2 Ma et al. (2016)	• Meta-analysis • 46 studies • Experimental, natural designs • Published & unpublished • Search range: From 1990	• K-6 • Countries: NR • Settings: NR • Sample size: NR	• Parental involvement • Control: NR • Duration: NR • Population Focus: NR	• Academic (language, mathematics, science, and others)	<u>Correlational study:</u> • learning outcomes and parental involvement related (d=0.51*, p<0.05); • Types of involvement significantly related to learning outcomes: home discussion, home supervision, home-school connection, and school participation, behavioural involvement, intellectual involvement • Grade level moderates • Follow up: NR	+ (68%)	Pr: 100% EYS: 72%	Supported (RQ1 only)
#486 Patall et al (2008)	• Meta-analysis • 14 experimental studies + cross-sectional • 6 RCT, 3CRCT, 5 QED • Published & unpublished • Search range: 1987-2004	• Grade: K-12 • Countries: USA & Canada • Settings: NR • Sample size: 1264 (IG: 642 vs CG: 622)	• Parental training for homework involvement • Control: No parent training • Duration: 4-35 weeks • Population Focus: Universal	• Student achievement.	• RCT: Academic achievement d=0.09, ns, k=6; homework competition d=0.28[d], k=4; problems with homework d= -1.20[*], k=3 • CRCT: d=0.01, ns, k=3 • QED: d=0.22*,k=5 <u>Grade (RCT only)</u> Grade 2-5: d=0.23*; p<0.05, k=3 Grade 6-8: d=-0.18*; k=2 <u>EYS:</u> EYS: No pooled ES, 3 relevant studies (mixed results).	+(63%)	Overall: 85% Pr 29% EYS RCT: 83% Pr, 16% EYS CRCT: 100% Pr, 33% EYS QED: 80% Pr, 40% EYS	Supported (RQ2)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
#116 See & Gorard (2015)	• Systematic review • 1008 (k=77 about parental involvement) • RCTs, QES, observational, regression discontinuity • Published & unpublished • Search range: Jan 2001-Jan2011	• Pre-school to school and beyond • Countries: NR • Settings: NR • Sample size: NR	• Any parental involvement strategies for child's behaviour or formal education • Control: NR • Duration: NR • Population Focus: NR	• School readiness, • Attendance, • School adjustment, • Likelihood of staying and exclusion from school, • Cognitive • Post-compulsory education participation	• For PreK children +ve association for parent reading to child with school readiness and success • For school-aged children, home-school partnership and parental interest in children's academic activities positively associated with school outcomes Follow up: NR	-(46%)	NR	Unknown (RQ2)
#513 Sénéchal & Young (2008)	• Meta-analysis • 16 studies • 12 experimental, 4 QED • Published • Search range: NR	• Grade: K-3 • Countries: Hong Kong, US, England, Australia, Canada • Settings: NR • Sample size: 1340 families	• Parent child reading activities • Control: NR • Duration: 1-13.5h, 0.5-36mo. • Population Focus: Universal and targeted	• Student literacy[e]	• Overall: d=0.65*, k=16 • No difference between inclusion of supportive feedback (d=0.62*, k=10) and no supportive feedback (d= 0.70*, k=6) • Early literacy d= 0.46*, k=3; word reading d= 0.31, ns, k=3; reading comprehension d=0.46*, k=3; composite measure d= 0.69*, k=6 • Kindergarten d= 0.51*, k=5; grade 1-3 d=0.74*, k=11.	++(83%)	EYS: 100%	Supported (RQ2)
#87 Semke & Sheridan (2012)	• Systematic review • 18 studies • Mostly descriptive (1 Experimental, 2 QES) • Published (peer-reviewed journals) • Search range: 1995-2010	• Grade: K-12 • Country: North America • Settings: Rural • Sample size: NR	• family involvement, family-school partnerships, and school-community partnerships • Control: NR • Duration: NR • Population Focus: NR	• Academic • social-emotional, • behavioural	• Identified the importance of family-school connections in rural areas as a common theme; • Participation in home-school connection programs considered beneficial in several studies • Many studies emphasized the importance of school-community connection • Barriers to parent-school connections included geographic	+ (58%)	Pr: 39% EYS: 33%	Unknown (RQ2)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
					isolation, poverty, family conditions and fun			
#618 Ttofi et al (2011)	• Systematic review and meta-analysis • 44 program evaluations (17 rel.) • 14 RCT, 21 QED,9 age cohort design (• Published & unpublished • Search range: 1983-2009	• K-high school • Countries: 15 USA/Canada vs. 29 others (25 Europe) • Settings: NR • Sample size: NR	• Anti-bullying programs • Control: No anti-bullying program • Duration: ≤40 days (23 studies) vs. ≥270 days (20 studies) • Population focus: NR	• Bullying and victimisation	Overall Odds Ratio: • bullying 1.36*, k = 38, • victimisation 1.29*, k=38. Programs which included parent training/meetings: • bully prevention (OR=1.57, k=17) • victimisation (OR=1.41, k=17)	+ (57.4%)	Pr: 43%<10 years EYS: NR	Supported

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d; d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES= Effect size; EYS=Early Years of School; #ES=number of ES; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC= Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns=not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; ~=approximately; * 95% CI **does not** encompass zero; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Shared reading, emphasised partnership, checking homework, communication between parents and teachers, head start program, ESL teaching; [d]p<0.10; [e] Word reading, early literacy skills, reading comprehension

Domain 13: Community-school partnerships

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
#602 Celio et al (2011)	• Meta-analysis • 62 studies • 19 RCT, 43 QED • Published & unpublished • Search range: 1970-2008	• Elementary-professional school e.g. medical school • Countries: NR • Settings: NR • n= 11,837	• Service learning (integration of community service with academic curriculum) • Control: NR • Duration: NR • Population Focus: NR	• Attitudes toward self • Attitudes toward school and learning • Civic engagement • Social skills • Academic achievement	• Overall g= 0.28*,p<0.05, k=62. • Attitudes toward self (g: 0.28*, p<0.05, k=35); attitudes toward school and learning (g: 0.28*, p<0.05, k=12); civic engagement (g: 0.27*, p<0.05, k=28); social skills (g: 0.30*, p<0.05, k=28); and academic achievement (g: 0.43*, p<0.05, k=17). Grade K-12 (g: 0.20*; p<0.05, k=19)	+(55.6%)	K-5: 5% EYS: NR	Unknown (RQ2)
#91 Davies (2013)	• Systematic review • 58 studies (9 relevant) • Qual & Quant • Published & unpublished • Search range: 2005-2011	• School age (5-18 years) • Countries: UK (esp. Scotland) • Settings: NR • Sample size: NR	• Environments & conditions that promote creative skill development • Control: NR • Duration: N/A • Population focus: NR	• Creativity • Academic • Emotional • Social skills • Motivation, • Engagement • Enthusiasm • Enjoyment, • Concentration • Attention	• 9 studies suggest involvement with outside agencies (e.g. community organisations and businesses) contribute to a creative learning environment Follow up: NR	+ (60%)	Pr: 41% (24/58) EYS: unclear (0 to 14%)	Unknown RQ1 only)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
#604 Henry et al (2017)	• QED (2 schools) • Number of teachers: NR	• Grade: 3-5 • Country: USA • Settings: Urban • n= 1290 (IG: 621 vs CG: 669) • Partners: School principal, school counsellor, other student service personnel, a parent, and faith-based volunteers	• Counsellor led, faith based, school family community (FBSFC) partnership called "Just Love"[c]. • Control: No FBSFC • Dose/duration: NR • Population Focus: Universal	• Student reading achievement (Florida Assessment for Instruction in Reading)	School level reading School x Time of Test interaction: • 2010-2011: ns • 2011-2012: p<0.05 • 2012-2013: p<0.01. Adopted vs non-adopted classrooms: School x Time interaction not significant in any year. Mentored vs selected non-mentored students: School x Time of Test: ns all three years	+(66%)	Pr: 100% EYS: NR	Preliminary (RQ2)
#612 Hoglund et al (2012)	• QED (17 schools) • Teachers: IG: 28–32 vs CG: 15– 17	• Grade 1 • Country: Canada • Setting: Public school • n=432 (IG: 290 vs CG: 142)	• Intervention vs Control • WITS Primary Program [d] • Control: No WITS training • Population Focus: Universal	• Peer victimisation, help seeking, social-emotional adjustment	Linear changes in: • physical victimization (d=0.17*), • relational victimisation (d=0.20*), • social competence (d=0.20*), • physical aggression (d=0.09*) • help seeking and internalising (d= 0.04 and 0.10 respectively – both ns)	++84%	EYS: 100%	Preliminary (RQ2)
#615 Lee & Stuart (2013)	• QED (20 schools) • Number of teachers: NR	• Grade 3, 5 & 7 • Country: Australia (QLD) • Setting: Catholic schools • n= 2758 (IG: 1526 vs CG: 1232)	• HPS- Health Promoting School; • partnership strategies: constant communication, shared visions, staff empowerment[e] • Control: No HPS	• Student resilience and protective factors.	Positive program effects : • Student family connection p<0.001 • Students community connection score p<0.001 • Student peer score p=0.011 • Student resilience p<0.001	+68%	NR Mean age 10 years	Preliminary (RQ2)

# Author (date)	Study characteristics	Sample	Intervention or Concept	Outcomes	Main Findings[a]	Quality Rating	Overall Relevance[b]	Evidence Rating
			• Dose/duration: 18 months • Population Focus: Universal		Age had a –ve effect on protective effect of HPS on resilience: (B= -0.132, beta = -0.072, SE =0.038, p=0.001), Follow up: 18months			
#617 Weaver et al (2017)	• QED (3 schools) • Teachers: 15 (12 classroom + 3 Phys Ed) teachers	• Grade 1-3 • Country: USA • Setting: NR • n: 222 (IG: 175 vs. CG: 47) • Implementer: PE/classroom teachers	• Partnerships for Active Children in Elementary Schools (PACES) • Control: routine practice • Dose/duration: NR (movement integration activities last 5-10min) • Population Focus: Universal	• Physical Activity	• All day: % children accumulating 30 min MVPA girls ES 0.65 (ns) and boys ES 0.38 (ns); % time in MVPA boys ES 2.13* and girls ES 0.70*; % time sedentary activity boys ES 0.85 (ns) and girls ES -0.19 (ns). • Class-time: % time MVPA boys ES 2.12 (ns), girls ES 0.82 (ns); % time sedentary activity boys ES 0.92 (ns), girls -0.04 (ns). • Physical education time: % time MVPA boys ES 0.73* and girls 0.56 (ns); % time sedentary activity boys ES 0.59 (ns) and girls ES 0.11 (ns).	+68%	EYS: 100%	Preliminary (RQ2)

CG=Control group; CI=Confidence Interval; CRCT=Cluster Randomised Control Trial; d=Cohens d; d'= Procedures described by Lipsey and Wilson (2001) and Sedlmeier and Gigerenzer (1989) were used to estimate effect sizes when unadjusted standard deviations were not available; ES= Effect size; #ES=number of ES; EYS=Early Years of School; g=hedges g; IG=Intervention group; k=number of studies; K=Kinder; MA=Meta-analysis; MC=Matched Control/Comparison; MS=Meta-synthesis; n=sample size; NR=Not Reported; ns= not statistically significant; PD= Professional development; PP=Pre-Post; Pr= Primary level studies; QES=Quasi-experimental studies; Qual= Qualitative studies; Quant=Quantitative studies; RCT=Randomised controlled trial; SR=Systematic review; TR=Teacher-rated; Tx=Treatment; WWC=What Works Clearinghouse; ~=approximately; * 95% CI **does not** encompass zero; [a] Where fixed and random effects were presented, results for random effects models have been extracted; [b] Proportion relevant was calculated based on the number of studies (or cohorts or effect sizes) where participants were exclusively within the target grade range (i.e. K-6 for primary students, K-3 for EYS or average age from 4 to 8 years). The proportion relevant in this table is for the meta-analysis overall, not for specific identified strategies; [c] Just Love is a counsellor-led, faith-based, school-family partnership including student mentoring, volunteer teaching aide, and school-wide incentive and enrichment program; [d] WITS is a Community based, whole school peer victimisation prevention program); [e] Schools were connected with the local communities and various organizations such as local city council, local Departments of Health and Education, and NGOs, which provided the school with a range of support services and resources

Appendix K: Proportion of studies with low SES samples

Reported low SES information for Well Supported and Supported strategies rated Very Plausible or Plausible to apply to the early years of school

Relevant domains and strategies: (i.e. Well Supported and Supported strategies rated Very Plausible or Plausible to apply to the early years of school identified in meta-analyses or systematic reviews reporting the proportion of studies or effects with predominantly low SES samples)	**Support & Plausibility ratings**	**% Studies/ effects with predominantly low SES samples**
Curriculum content and delivery		
Balanced reading and writing instruction (Graham 2018)	WS, VP	<25%
Inference instruction for reading comprehension (Elleman 2017)	S, P	15%
Phonemic awareness, phonics, fluency, and reading comprehension (Hammill 2006/ Ehri et al 2001)	S, VP	24%
Peer teaching and co-operative learning approaches		
Targeted peer tutoring in literacy and numeracy (Kunsch 2007)	S, P	41%
Peer tutoring across a range of student abilities (Leung 2015)	S, P	17%
Peer tutoring across a range of student abilities (Zeneli 2016)	S, P	9%
Peer Assisted Learning interventions across a range of subject areas (Rohbeck 2003)	S, VP	26%
Technology-assisted teaching and learning		
Technology-enhanced stories (Takacs 2015)	WS, P	42%
Various technology for reading instruction (Cheung 2012)	WS, VP	**80%**
Social-emotional and behavioural (SEB) interventions to promote a positive school climate		
Teacher training in classroom management programs and strategies (Korpershoek et al 2016)	WS, VP	**50%**
Mindfulness interventions (Zenner 2014)	S, P	25%
Specific/universal SEB programs (Sklad 2012)	S, P	33%
Staff-student relationships		
Correlation only: Teacher-student interaction and student cognitive functioning (Vandenbrouke 2018)	WS, VP	**64%**
Partnerships with families		
Parent-Child Reading (Senechal 2008)	WS, VP	31%

WS=Well Supported; S=Supported; P=Plausible, VP=Very Plausible

Appendix L: Definitions for school quality indicators

Balanced reading and writing: Literacy instruction whereby no more than 60% of time is allocated to either reading or writing

Best practice teacher coaching : To be considered best practice, coaching should be characterised by at least four of the six following criteria: individualised (1:1 feedback), intensive (conducted at least fortnightly), sustained (provided over a substantive period of time), context-specific (tailored to the teachers class), focussed (provides specific tasks for teachers to practice), and combined with curriculum-specific materials/resources

Classroom management strategies: The strategies teachers use in the classroom to create an environment that supports and facilitates student learning. Examples of evidence-based classroom management strategies include PATHS, the Good Behaviour Game, the Incredible Years Teacher Classroom Management Program, Proactive Classroom Management Program

Classroom teachers: Teaching staff who regularly supervise the main literacy and numeracy instructional blocks (i.e. not casual relief teachers or specialist subject teachers such as those delivering instruction in Art, Science, Technology, Physical Education, or Languages Other Than English for example)

Differentiated teaching: Modifications to instructional delivery that enable teachers to tailor instruction to the needs of students across a range of abilities and learning needs

Evidence-based interventions: Strategies that have demonstrated positive and statistically significant effects of at least moderate magnitude (i.e. standardised mean differences of 0.3 or more) or practical importance in at least two randomised controlled trials, on relevant outcomes (i.e. student academic performance or psychosocial development).

Formal training: Participation in external professional development opportunities (such as workshops run by independent organisations)

Informal professional development: Training or skill development opportunities that are developed and implemented internally by schools (e.g. coaching from more senior teachers in same school) or between school clusters (e.g. communities of professional development meetings involving teachers from multiple schools sharing knowledge or experience or 'practice wisdom').

K-3: The first year of school to grade 3 (children are approximately 5 to 8 years of age)*.

Manipulatives: Objects with which students physically interact to aid learning. Examples of manipulatives include play money, counting blocks, and cuisine rods (not use of rulers, scales or calculators).

Materials to support reading at home: Examples include materials that describe dialogic reading practices; interactive listening to child read; tutoring specific skills such as alphabet knowledge or word reading strategies

Peer tutoring: Structured activities in which same-age, cross-ability, student pairs receive explicit instruction and guidance in tutoring one another.

Tier 2 intervention: Additional small-group instruction for students who do not make adequate progress with classroom instruction or who fail to meet benchmarks on screening measures (intensity of intervention is varied according to group size, frequency and duration of intervention, and level of provider training).

Tier 3 intervention: intensive one-to-one supports specifically targeting skills deficits that are provided when students do not adequately respond to Tier 1 or Tier 2 instruction.

Safety at school: Is defined in terms of response to school survey item "I feel safe at school"[^], or items assessing whether students have experienced bullying or physical or verbal maltreatment (e.g. "I have been bullied at my school this term", "I have often been teased in an unpleasant way or called names at my school"[✦])

School encouragement of parent-child reading: Is defined in terms of response to school survey item, "This school works with me to support my child's learning"[▲], or other similar items available through state school surveys (e.g "This school works with me to support my child's learning", "Staff at this school are responsive to my enquiries"[✢])

Small group: Groups comprising no more than six students

*As the term 'kinder' has often been used to refer to the first year of formal schooling (both in the international literature and some Australian states) we use the terms K-3 to refer to the early years of school.

[▲] An agreed student item in the Australian Curriculum and Assessment Reporting Authority (ACARA) School Survey

[✦] Survey item examples from the Framework for Improving Student Outcomes, Student Attitudes Survey

[✢] Survey item examples from Queensland School Opinion Survey, Parent Items

ACKNOWLEDGEMENTS

Restacking the Odds wishes to acknowledge the contributions made by the following individuals in the preparation of this report. Dr John Ainley (Australian Council for Educational Research) provided expert guidance and feedback on the indicators developed to assess school quality. Phoebe Quinn (MCRI) and Shuaijun Guo (MCRI) assisted with development of the literature search and data extraction.