



Education  
Endowment  
Foundation

# Improving Behaviour in Schools: Evidence Review

December 2019

Darren Moore<sup>1</sup>  
Simon Benham-Clarke<sup>2</sup>  
Ralphy Kenchington<sup>2</sup>  
Chris Boyle<sup>1</sup>  
Tamsin Ford<sup>2</sup>  
Rachel Hayes<sup>2</sup>  
Morwenna Rogers<sup>2</sup>

<sup>1</sup> Graduate School of Education, University of Exeter

<sup>2</sup> University of Exeter Medical School



Please cite this report as follows: Moore, D., Benham-Clarke, S., Kenchington, R., Boyle, C., Ford, T., Hayes, R. and Rogers, M. (2019). *Improving Behaviour in Schools: Evidence Review*. London: Education Endowment Foundation. The report is available from:

[https://educationendowmentfoundation.org.uk/public/files/Improving\\_Behaviour\\_in\\_Schools\\_Evidence\\_Review.pdf](https://educationendowmentfoundation.org.uk/public/files/Improving_Behaviour_in_Schools_Evidence_Review.pdf)

The EEF Guidance Report *Improving Behaviour in Schools* is available at:

<https://educationendowmentfoundation.org.uk/tools/guidance-reports/improving-behaviour-in-schools/>

### **Acknowledgements**

With thanks to Naomi Shaw for assistance with screening and Lucy Goble for assistance with data extraction. Also to the helpful comments from the guidance panel: Neil Humphrey, Alex Sutherland, Jane Bateman, John D'abbro, Jenny Thompson, Michelle Long.

## **Contents**

<b>Background</b>	<b>3</b>
<b>Aims</b>	<b>9</b>
<b>Review 1 Method</b>	<b>10</b>
<b>Review 1 Findings</b>	<b>12</b>
<b>Review 2 and 3 Methods</b>	<b>32</b>
<b>Review 2 and 3 Findings</b>	<b>37</b>
<b>QCA Method</b>	<b>90</b>
<b>QCA Findings</b>	<b>92</b>
<b>Discussion</b>	<b>96</b>
<b>References</b>	<b>101</b>
<b>Appendices</b>	<b>110</b>

## Background

Behaviour in schools and classroom management has been the focus of a great deal of research, theory, policy and media attention (Ball et al., 2012). Despite this, pupil behaviour remains a challenging area for all stakeholders in the education of children and young people and is commonly cited as one of the most difficult tasks that both experienced and new teachers have to contend with in schools (Barmby, 2006; Jennings & Greenberg, 2009; Kokkinos, 2007). Research suggests that both (i) understanding of pupil behaviour and (ii) how best to train and support teachers to manage pupil behaviour is contested (Beaman et al., 2007; Powell & Tod, 2004). Todd & Ellis (2018) highlight that many in education have a view on what behaviour management is, how it should be approached and a conviction that more can be done. Although there is certainly no consensus across the sector, support can be found in a plethora of guidance and advice resources, from websites to training consultants. This support can over simplify the complexity in achieving a complete understanding of behavioural influences, behaviour management and the evidence base needed in the field.

Defining behaviour is not straightforward and there are many alternative definitions (Department for Education, 2012). Behaviour at its broadest can relate to any action that schoolchildren take and therefore could relate to choices about relationships with peers, eating and physical activity. While important, such behaviours are outside of the remit of this review, which is focused more on understanding school and classroom behaviour that affects learning and either meets or challenges the expectations for pupil conduct at school. However, “behaviour” in the context of this review does not only refer to poor behaviour or misbehaviour, such as Cameron’s (1998) classification that specified the following categories; aggressive behaviour, physically disruptive behaviour, socially disruptive behaviour, authority-challenging behaviour and self-disruptive behaviour. In this review, we also include positive behaviour for learning (Ellis & Tod, 2018), such as concentration, prosocial behaviour and engagement are relevant. So our definition of behaviour includes both negative and positive actions that are open to subjective interpretations, even to the extent a specific behaviour may be of concern to one teacher but not to another, or of concern in the classroom but not on the playground (Watkins & Wagner, 2000).

Difficulties in classroom management often lead to stress, burnout and exit from the teaching profession (Aloe et al., 2014), as well as being a deterrent for those considering teaching as a career (Day et al., 2006; DfE, 2010; Jacobson, 2016; Ozdemir, 2007). It is also cited as a challenge for headteachers across all school phases (Leithwood & Day, 2008). Ineffective classroom management can lead to pupil disengagement, aggression, low attendance and

bullying (DfE, 2010; Zyngier, 2007). There is a need not just for interventions and strategies that target the severe behaviour concerns that arise from a minority of students, but also the low level disruption that can be a concern across classes to the point that it may undermine learning (Ofsted, 2014). However, recent data from Ofsted indicates that behaviour in schools is generally good with Ofsted giving 85% of all schools overall good or outstanding ratings (Ofsted, 2019a). Also, from a teacher voice survey (Department of Education, 2018) 86 per cent of senior leaders rated behaviour as good or very good, yet this reduced to 59 per cent for classroom teachers. The behaviour from the majority of children and young people in the majority of schools is therefore good.

Schools cannot function well if pupils are frequently absent or do not feel safe. While attendance at school and bullying are not wholly the preserve of schools (DfE, 2015), schools have an important role to play, this is of particular concern as ‘... both authorised and unauthorised absence rates have increased since last year, the rate of the latter now being the highest since records began’ (DfE, 2019, p. 1). While there is a lack of clear statistics reporting the incidence of bullying in schools, NSPCC report that they have a call from a child on average every 25 seconds and 1 in 5 children have suffered abuse or neglect (NSPCC, 2018). Increasing school absence (DfE, 2016) or being bullied (Brown, 2018) are linked to lower attainment. For a focus on preventing bullying see the systematic review and meta-analysis from Ttofi and Farrington (2009). Hence, school approaches to prevent and respond to absences and bullying are likely to involve the creation and sustained high quality behaviour management throughout the whole school with support from parents and other stakeholders.

In light of concerns regarding increases in school exclusions, and that despite the same school exclusion framework applying to all state funded schools in England, there are variations ‘... in exclusion rates between schools, areas of the country, and pupils with different characteristics’ (DfE, 2018a, p.1), the recently published Timpson Review commissioned by the DfE reviewed school exclusions (DfE, 2018a) and also considers the practice of off-rolling (Timpson, 2019). Of more relevance here, this review will examine the ‘practice in schools in relation to behaviour management and exclusions. This includes identifying effective approaches which improve outcomes, particularly for those groups disproportionately likely to be excluded.’ (p. 2).

Tom Bennett’s independent review of behaviour in schools focusses on the approaches school leaders can take to develop the culture in their schools to ‘promote excellent behaviour’ (Bennett, 2017, p. 30). With reference to examples and case studies he asserts

that leaders ought to focus on whole school culture to benefit behaviour, rather than focussing on teachers in isolation. These approaches are presented as recommendations that ‘... reflect the three stages of promoting a school culture that deliberately and carefully optimises conduct, character and academic achievement - designing the culture, building the culture in detail and maintaining the culture.’ (p.30). He concludes with cautionary advice that schools need to adhere to the recommendations clearly, consistently, and realistically across the school, but nonetheless with high expectations of what can be achieved. There is recognition of a variety of obstacles to achieving this adherence, so the review offers some suggestions for school leaders. In addition, Ofsted are reviewing their framework for inspection to focus separately on firstly school’s management of behaviour and pupil attitudes, and secondly on personal development. Behaviour and attitudes incorporates related aspects including study skills, resilience and relationships across the school body (Ofsted, 2019b).

How “behaviour” has been viewed in the above and other policy related documents has changed and this may influence the actions of schools. Reports such as ‘Pupil Behaviour in schools in England’ (Department of Education, 2012) and ‘Below the radar’ (Ofsted, 2014) signified ‘official’ concerns about the impact of misbehaviour on attainment and wellbeing. These concerns have manifested themselves in different ways, for example, more emphasis on discipline, although the new Common Inspection Framework distinguishes between behaviour and discipline and pupils’ wider personal development (Ofsted, 2019b). More recently off-rolling (Timpson, 2019) and mental health and wellbeing (Department of Education, 2018b) have all had increased focus. As these are commonly linked to behaviour in schools, there has been increased guidance and regulations regarding the approach and actions schools could and must take to manage behaviours.

Both reactive and preventative interventions may improve behaviour, while a reduction in challenging behaviour can also help to improve other variables such as classroom climate, attendance and attainment (Gastic, 2008; Marzano & Marzano, 2003). Previous reviews of research that investigated the effectiveness of classroom management or whole school approaches have been conducted but often focus on a particular type of intervention or outcome (Korpershoek et al., 2016; Oliver et al., 2011; Whear et al., 2013; Wilson & Lipsey, 2007). Models focusing on explaining pupil behaviour need updating and clearer links to the evidence they use to explain school behaviour (Powell & Tod, 2004; Scheuermann & Hall, 2015). Such models are often used in teacher training texts (e.g. Cooper & Elton-Chalcraft,

2018; Glazzard et al., 2014; Graham-Matheson, 2014; James, 2016), so there is scope for this review to contribute to teachers' understanding of behaviour.

Teachers want their pupils to learn, hence behaviours that are disruptive to a child's own learning or that of others in their class can be viewed as behaviours to 'manage'. However, a shift of focus from managing a child's behaviour and towards teaching a child learning behaviours may be beneficial (Bitsika, 2003; Ellis & Tod, 2018; McDermott et al, 2001; Nolan, 2011; Norwich & Rovoli, 1993; Núñez and León, 2015; Powell & Tod, 2004). While accepting that teachers will need to manage behaviour at times, promoting learning behaviours could be seen as not only in the interest of the child and their peers in class but also of the whole school, as well as the child throughout their education and adulthood. Moreover, a focus on teaching learning behaviours seems to fit well with the role of a teacher and be in their sphere of control, whereas managing a child's misbehaviour may be complex and challenging. Furthermore, even if a teacher successfully manages a child's behaviour this does not necessarily lead to that child learning. The Behaviour for Learning conceptual framework (see figure 1) identifies three core pupil relationships, those with self, others, and the curriculum. Each relation impacts on the other, and positive change can be achieved by recognising which of these relationships need to be developed or strengthened (Ellis & Tod, 2018).

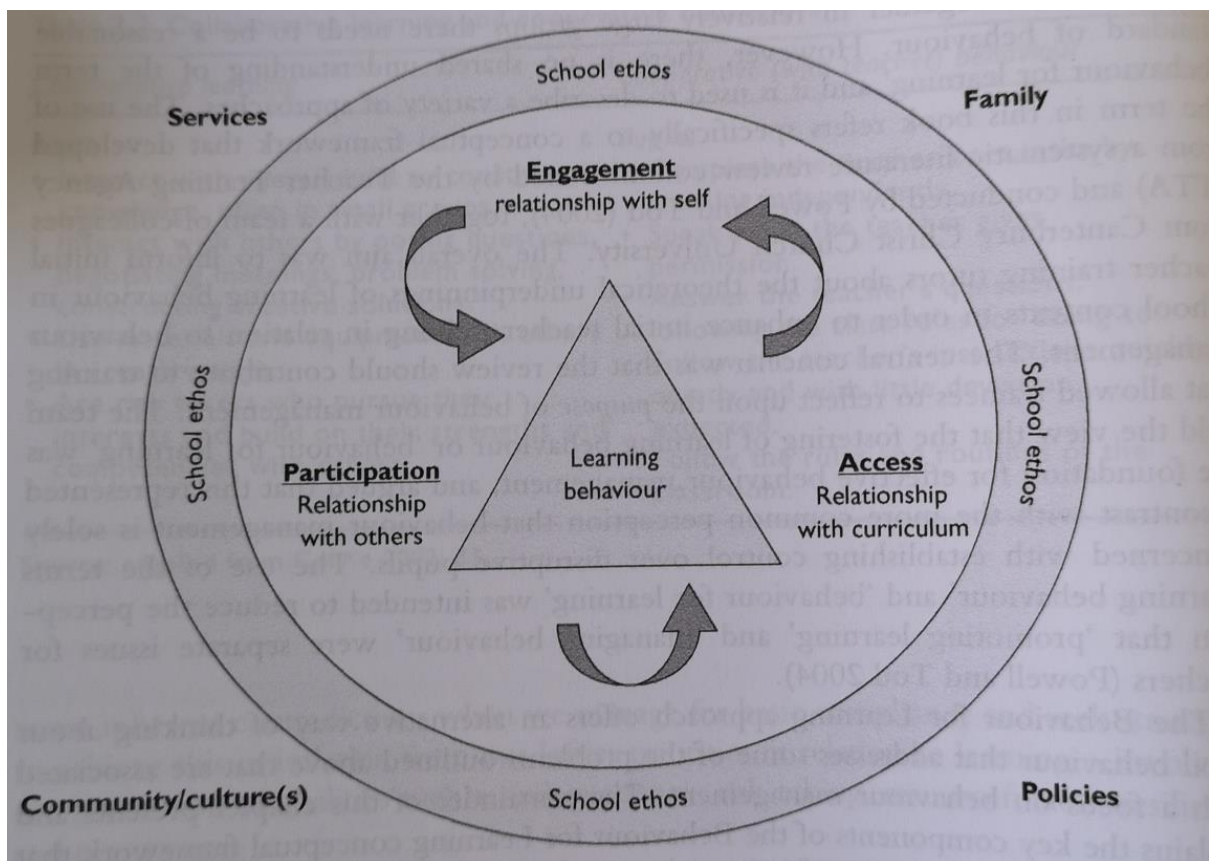


Figure 1. The behaviour for learning conceptual framework, from Ellis & Tod (2018, p. 38).

While there are a variety of interventions that teachers can focus on at a class level and members of school staff can be trained to target undesirable behaviours, school behaviour approaches can also involve consistency and coherence at a whole school level. These often relate to promoting inclusion, school culture, positive role models and organisational principles (Adolphus et al, 2013; Bodin, 2016; Bradshaw, 2012; Garner, 2011; Hershfeldt et al, 2009; Smith, 2010). Public Health England (2015) provide a model of elements of a whole school approach for wellbeing interventions, this may be of relevance in considering whole school approaches to behaviour (see figure 2).

However, there are a very wide range of interventions that may theoretically improve aspects of school behaviour. Reviewing all of these would be a prohibitively large task and would duplicate existing high quality reviews (e.g. Bruhn et al., 2015; Daly-Smith et al., 2018; Evans et al., 2003; Flower et al., 2014; Korpershoek et al., 2016; Law et al., 2012; Losinski et al., 2014; Machalicek et al., 2007; Maggin et al., 2011; Maggin et al., 2012; McKenna et al., 2016; Moore et al., 2018; Whear et al., 2013; Wilson & Lipsey, 2007), including those already covered under other areas of EEF focus – examples include social and emotional learning, parent engagement, thinking skills and self-regulation, physical activity, social



skills, interventions to improve symptoms of externalising disorders. See

<https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/>



Figure 2. Principles to promoting a whole school and college approach to emotional health and wellbeing. From PHE (2015, p. 6).

## Aims

This evidence review synthesises the best available international evidence regarding approaches to behaviour in schools to:

- Produce an overarching model or framework explaining why school pupils may misbehave (Review 1)
- Review the effectiveness of classroom-based approaches to behaviour trialled in robust research studies (Review 2)
- Review the effectiveness of school-wide approaches to behaviour and identify the gaps for robust research studies to explore frequently used strategies (Review 3)

- Analyse what components of the universal behaviour interventions predict improved behaviour outcomes (Qualitative comparative analysis)

## Review 1 Method

To provide an overview of evidence on why pupils may misbehave in schools and produce a model summarising this evidence, we used the following method:

### Inclusion criteria

We anticipated that the literature included for this review would be unlikely to include intervention studies. Therefore, the following inclusion criteria were specified:

*Article focus:* Behaviour of children and young people in school settings. Articles needed to include a written or diagrammatic model, framework or explanation that can help to explain why school students misbehave.

*Study design/Publication type:* Any. We were interested in recording the evidence used to inform the frameworks and therefore prioritised review evidence and relevant primary and secondary analysis with similar research questions as this review.

*Date:* Any

*Language:* English only.

*Country:* At full text screening we would have excluded any frameworks from countries with very different school systems compared to England.

### Search strategy

On 6<sup>th</sup> and 7<sup>th</sup> November 2018 the following databases were searched:

ERIC, Education Research Complete and the British Education Index (via EBSCOhost), the Australian Education Index and ASSIA (via ProQuest), PsycINFO (via OvidSp), Social Science Citation Index (via Web of Science), the Cochrane Library and the Campbell Library.

The database searches were designed and run by an information specialist (MR). The search combined terms for frameworks/models, behaviour and school. An example search strategy can be seen in appendix 1.

Forwards and backwards citation chasing was performed for one key review of literature (Powell & Tod, 2004).

We also searched the following websites for potentially unpublished literature: Education Endowment Foundation; What Works Clearing house; Department for Education; The Schools, Students and Teachers Network; Devon LEA; Parentkind.

Key journals checked were those where multiple included studies were published.

### Study selection

References located by the search were uploaded to reference management software (Endnote X8) and duplicate studies were removed. Two reviewers screened the titles and abstracts, after which DM independently screened all records that were not excluded. Full texts were retrieved for those records included at this point. One reviewer initially screened all of these and DM independently screened all decisions that were unclear. A PRISMA-style flowchart was produced with the reason for exclusion of each full-text article reported (Moher et al., 2009).

### Data extraction and quality appraisal

We extracted study details from included studies including country, design, sample, school level; as well as details of the explanatory framework of behaviour detailed in the study. We extracted any further information regarding the evidence for the framework (particularly for non-primary research studies) as this was the marker of quality appraisal and provided a rating as to how evidence-informed the frameworks from studies were. This allowed an assessment of evidence for potential explanatory factors for misbehaviour that appear in the final model produced from this review. This meant that quality appraisal could be conducted at the level of synthesis, by considering the strength and robustness of evidence for each explanatory factor linked to behaviour that appeared in the synthesised model. This fits quality appraisal recommendations for systematic reviews of intervention effectiveness that are interested in several different outcomes (e.g. GRADE, Guyatt et al., 2008). Quality appraisal at the level of included study was not considered as there was anticipated to be multiple studies providing evidence for some explanatory factors.

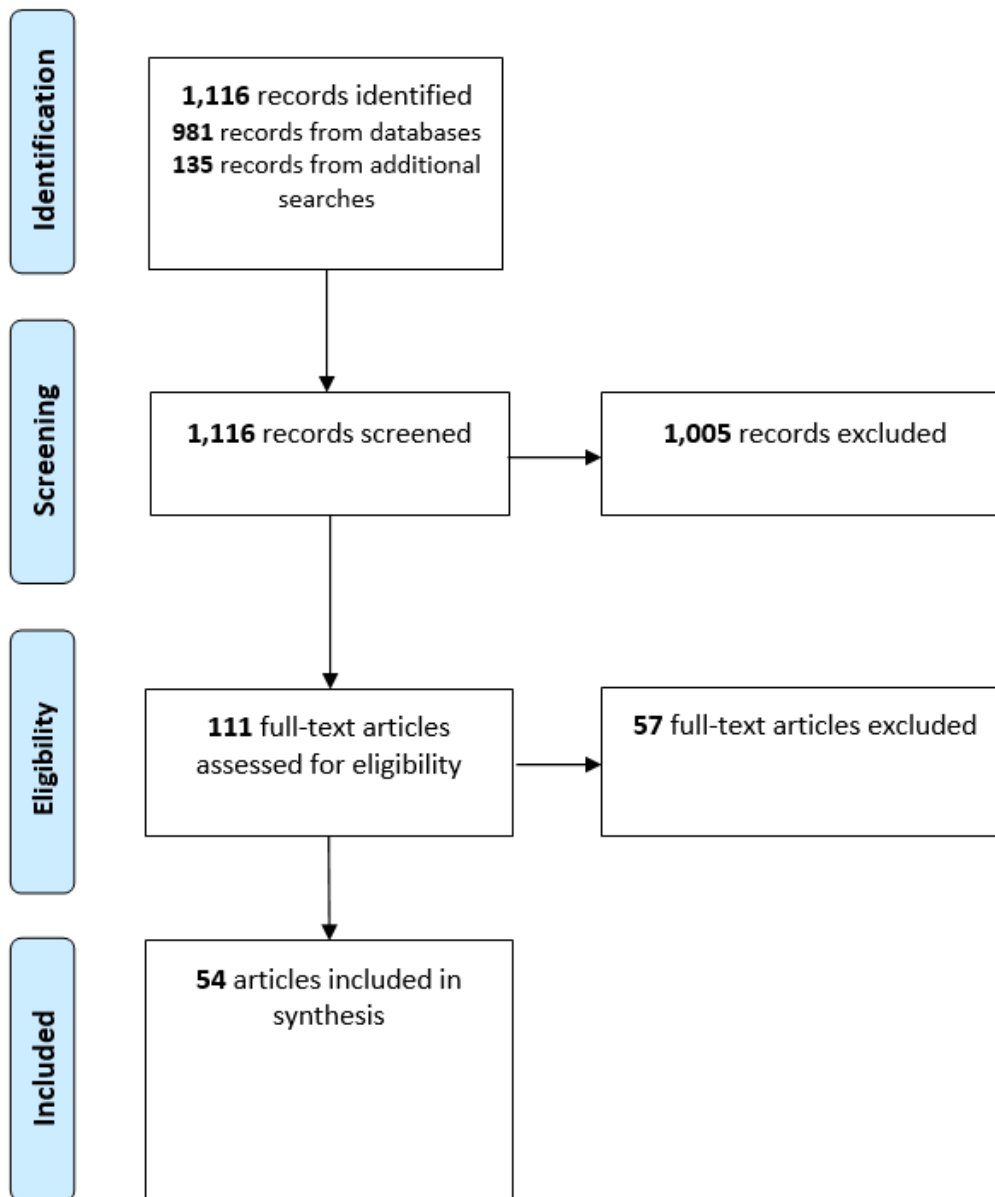
### Synthesis

The synthesis of included studies aimed to produce a model of explanatory factors that may inform school behaviour management derived from the evidence reviewed while narrative synthesis was conducted to summarise and explain the final model and combine the literature that informs it. The model takes into account the strength of evidence that informs the factors identified. We consulted with the guidance panel as the model was drafted. Therefore, the organisation of explanatory factors into categories has seen a number of iterations.

## Review 1 Findings

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram in Figure 3 summarises the process of study selection. Initial database searching gave 981 records to title and abstract screen. Seventy-seven of these records were considered relevant and their full texts were retrieved and screened, leading to 45 included studies. Nine more studies were later included after citation chasing for the Powell and Tod (2004) review. The majority of studies excluded at full text were focusing on topics other than school behaviour, e.g. violence, bullying, achievement.

Figure 3. PRISMA flow diagram showing study selection process for Review 1



### Descriptive synthesis

The summary details of the included studies can be seen in Table 1. These 54 studies were conducted from 1978 to 2018, mainly in the US (n=30), with only nine from the UK. Primary data was used exclusively in the majority of studies (n=30 studies), a combination of primary and secondary in eight studies, solely secondary data in two studies and the remainder (n=14) were review or discussion based. Primary school level participants were the focus in 14 studies, secondary participants in 12 studies and two studies had an explicit focus on special schools.

For those studies that reported the sample size, child participants were the focus in the majority of studies (n=36), with a range from sample sizes (1 to 5126) Teachers were included in 19 studies and parents or carers in nine studies. Eight studies focused on all children in an educational setting. Gender was reported in most studies (n=35) with one study exclusively focussing on boys. We made a note of where the included studies drew on existing models and frameworks in their work, the expectancy-value theory of motivation (Wigfield & Eccles 2000) was used as a guiding framework in three studies, but otherwise there was little overlap in theoretical frameworks used. In terms of quality appraisal, Appendix 2 shows the quality rating of each study and how they map onto each explanatory factor. There is typically good quality evidence for the majority of the explanatory factors that schools can manage directly. The weaker quality evidence of the link between school behaviour and out of school behaviour, ability and culture suggests less dependability in these explanatory factors according to evidence located.

Table 1. Description of included studies in Review 1.

<b>Study details</b>	<b>Country</b>	<b>Design</b>	<b>Model</b>	<b>Findings</b>	<b>Sample size</b>	<b>School level</b>	<b>% of females</b>	<b>Age M(SD)</b>	<b>Type of sample</b>
Adolphus et al. (2013)	UK	Literature review	School breakfast program	Breakfast improves on-task behaviours	Various	All	NR	NR	Special and non-special schools
Bear & Rys (1994)	US	Primary via interviews and questionnaires	Moral development and socio-economic status	There is a reciprocal relationship between maladjustment behaviours and socio-economic status	133 children	Primary	45%	NR	Various levels of adjustment.
Bidell & Deacon (2010)	US	Primary via interviews and questionnaires	School counselling	Low self-concept students exhibit disruptive classroom behaviours (DCB)	92 children	Secondary	72%	16.26(.09)	School counsellor selected students with and without DCB
Bitsika (2003)	Australia	Discussion with primary data via case studies.	Functional assessment framework	Meaningful and long-term changes to difficult behaviour result from teaching the student to behave differently	4 cases each with one child	All	NR	NR. An 8, 9, 10 and 14-year old child.	Those with ASD, ADHD, anger management problems, are disruptive or uncooperative
Boon (2011)	Australia	Primary data via school records and questionnaires	Interactions of school moves, coping and achievement	Positive coping strategies protect students from behavioural problems	1,050 children	Secondary	NR	NR - range 12 to 15 years	Three schools in North Queensland, of different socioeconomic status.
Borders et al (2004)	US	Primary data via questionnaires	Expectancy-value theory	Behavioural problems may result from a perceived lack of valued or feasible alternative behaviours	121 children and 4 teachers	Secondary	62%	16(1.22)	English teachers and their students in a multi-ethnic urban community.

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Cadieux (2003)	Canada	Primary data via a longitudinal study	Grade retention	Classroom behaviour is related to factors other than grade retention	67 children and their teachers	Primary	48%	In the first year 83 months (SD= 5.1)	Children from 3 school districts of the Outaouais region in the province of Quebec
Caughy et al (2007)	US	Primary data via interviews, observations and questionnaires. Secondary data via a census	Integrated Process Model	Lower levels of parental eliciting <sup>Note 1</sup> were associated with higher levels of internalizing and externalizing behaviour and lower levels of positive school adjustment	405 families	In or entering primary	49.9%	NR	Families from Baltimore City neighbourhoods
Cheung (1997)	Hong Kong	Primary data via a survey	Control and labelling theory	Peers deviant behaviour is a significantly better predictor of a child's deviant behaviour than media use, family, or school variables	1139 children	Secondary	56%	M = 14.8 Range 12 to 20 years	From schools randomly selected from urban areas, new towns and rural areas
Chirinos (2018)	US	Primary data via exam and questionnaire secondary data via PISA	Expectancy-value theory of motivation	Student perception of control and attributing academic success to effort can influence behaviours	2,488 children.	Secondary	50%	NR	From the 2012 U.S. PISA dataset., randomly selected from 162 public and private schools
Cornwall (2015)	UK	Discussion and opinion	Cognitive-behaviour and social construction - the 'human element' model	Understanding the human element underpins teaching challenging and vulnerable young people and improves the learner-teacher relationship generally	NR	Primary to secondary	NR	NR	Emphasis on those likely to be excluded



Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Daniels & Williams (1999)	UK	Descriptive paper	Framework for Intervention	Variations in any aspects of the behavioural environment can have a bearing on specific examples of behaviour	N/A	All	NR	NR	All
Dever (2016)	US	Primary data via questionnaires	Expectancy-value theory of motivation	Student motivation can be used to predict the early precursors of behavioural and emotional difficulties	5,126 children	Secondary	52.80%	NR. Range grades 9 to 12	One school district in a small city in South-East
Di Maggio et al (2016)	Italy	Primary data via a vocabulary test of children and teacher questionnaires	A mediation model	Targeting emotional knowledge and competencies may benefit social competence and prevent behavioural problems	240 children and their teachers	Kindergarten classes of preschools	51%	4.23(0.80)	All attending classes of schools in a big city of southern Italy.
Dinh et al (2001)	US	Secondary data via annual school surveys	Mediation model of acculturation and problem behaviour proneness	Parental involvement, but not self-esteem, plays a significant mediational role in children's problem behaviour proneness	330 children	Primary and middle schools	50%	NR. Range grades 4 to 8.	Hispanic students only.
DiStefano & Kamphaus (2008)	US	Primary data via questionnaires	Latent growth curve model of child behaviour change	Behavioural development follows a linear trajectory, although rates of decline vary in relation to perceived levels of risk	162 children and their teachers	Primary	49%	NR. Range 6 to 8 years.	A racially diverse sample from one school

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Ellis & Tod (2018)	UK	Opinion/views	Behaviour for learning	Schools should focus on promoting effective learning behaviours rather than stopping unwanted behaviours	N/A	All	N/A	NR	All
Eve (1978)	US	Article and exploratory primary data via questionnaires	Interrelationship of strain, culture conflict and social control theories	Social control has a greater containment effect on the motivation for deviant behaviour than adherence to community norms or as a response to frustration	300 children	Secondary	55.2%.	M = 17	All in South-Eastern city school.
Finn et al (2003)	US	Review of theory and research	Class sizes	Smaller class sizes have a positive impact on learning and social behaviours	N/A	Primary	N/A	NR	All
Garner (2011)	UK	Opinion on document, practitioner-knowledge and research	Role of leaders in the promotion of positive behaviour	Effective promotion of positive behaviour and learning, can be linked to leadership	NR	Primary and secondary	N/A	NR	All
Gottfredson et al (1994).	US	Primary data via questionnaires and secondary data via school records	Grade retention	Grade retention does not negatively affect problem behaviour, it does reduce rebellious behaviour and increases attachment to school	401 children	Middle	Promoted group 55% retained group 31%	Promoted group 12.41 (1.27) retained group 12.47 (1.10)	Those who scored in the bottom half of the achievement test in grades 6 and 7 in two southern, urban schools

<b>Study details</b>	<b>Country</b>	<b>Design</b>	<b>Model</b>	<b>Findings</b>	<b>Sample size</b>	<b>School level</b>	<b>% of females</b>	<b>Age M(SD)</b>	<b>Type of sample</b>
Güleç & Balçık (2011)	Turkey	Primary data via questionnaires	Teacher perspectives on undesirable behaviour	Undesirable behaviour results from family problems, and parental indifference to education', negative attitudes and behaviours, but not boring lessons, teacher's incapability or lack of teaching aids'	54 teachers	Primary	74.1%	NR	Five schools in Gölcük district of the Kocaeli province
Hastings (2005)	UK	A review of research literature	Actions in response to severe behaviour are escape or avoidance behaviours	Staff behaviour affects behaviour problems, and behaviour problems affect staff behaviour	NR	Special educational settings	N/A	N/A	Special education
Haynes (1990)	US	Primary data via questionnaires	Self-concept and adjustment	Self-concept dimensions and general classroom behaviour, group participation, and attitude toward authority are correlated	142 students 60 teachers	Middle School	Approx. 50%	NR. Range 10 to 13 years	Mainly African-American randomly selected children in 4 urban schools in North-East

<b>Study details</b>	<b>Country</b>	<b>Design</b>	<b>Model</b>	<b>Findings</b>	<b>Sample size</b>	<b>School level</b>	<b>% of females</b>	<b>Age M(SD)</b>	<b>Type of sample</b>
Heaven et al (2009)	Australia	Primary data via questionnaires over 4 years	A transactional model	There are bidirectional influences between psychoticism and adjustment and behavioural problems.	Teachers NR. Children 866	Secondary	Approx. 50%	Grade 8 13.63 (0.51) Grade 11 16.18 (0.46)	Those children retained or with problem behaviours
Hershfeldt et al (2009)	US	Primary data via a case study	Double-check: a cultural responsiveness framework	Double-check can improve the behaviour of culturally and linguistically diverse (CLD) children.	1 child	Primary	0%	NR	Culturally diverse students
Jackson & Frick (1998)	US	Primary via questionnaires	Compensatory, challenge and immunity/vulnerability models	Negative life events are not associated with adaptive or internalizing behaviour. Both negative life events and protective factors contribute externalizing behaviour. Girls with significant negative life events and social support demonstrate less internalizing behaviour	140 children	Primary to secondary	57%	NR. Range 8 to 13.6 years	A representative sample of children from several schools in a semi-rural southern town

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Kemp & Center (2003)	US	Primary via questionnaires	Eysenck's personality theory	Children with high psychoticism, extraversion and neuroticism traits are at greatest risk of developing conduct problems.	150 children	Middle, secondary and special day schools for severe emotional problems	22.70%	13.65 (1.9)	General education and those with emotional and behavioural difficulties matched on the variables of age, sex, and race.
Kerr et al (2000)	US	Primary via interviews and secondary via longitudinal data	Cumulative risk model	Children who have both a history of failure-to-thrive and maltreatment demonstrate more behaviour problems and worse school functioning than those who have neither	193 children and their teachers and parents	Primary	48%	M=73	Families recruited from paediatric clinics serving inner-city, low-income, primarily African-American families
Knutson et (2004)	US	Primary data via interviews, questionnaires and observations	Social disadvantage and neglectful parenting as precursors to antisocial and aggressive behaviour	Deficient parenting involving neglect contributes to the development of anti-social or delinquent behaviour and poor peer relations	671 children their teachers and parents	Primary	51%	First grade and fifth grade	High-risk neighbourhoods in a medium-sized metropolitan area with high rates of juvenile delinquency

<b>Study details</b>	<b>Country</b>	<b>Design</b>	<b>Model</b>	<b>Findings</b>	<b>Sample size</b>	<b>School level</b>	<b>% of females</b>	<b>Age M(SD)</b>	<b>Type of sample</b>
LaRoque (2008)	US	Primary data via interviews and questionnaires	Peer Group Processing Model	Peers and parents significantly affect student behavioural decisions. Students with EBD perceive less encouragement from school, parents, and peers	370 children	Middle and secondary	51% of those without disability, NR the rest	Grades 7 to 11	Schools around Wisconsin, a rural and suburban community, a medium-sized city, and an alternative school
Liao et al (2015)	US	Primary data via questionnaires and secondary longitudinal data	Disruptive behaviour in the wake of a trauma (school shooting)	Following a school shooting disruptive behaviours decreases over time	NR	Primary at baseline	52%	Most in grades 1 or 5	Participants of the Linking the Interests of Families and Teachers study most of whom were at schools within a 15-mile radius of a prior school shooting
Liasidou (2016)	Cyprus	Theoretical paper	Intersectionality-based policy analysis (IBPA) framework	Understanding and management students' problem behaviour needs a holistic and socially just approach	N/A	All	N/A	N/A	Children with social, emotional and behaviour difficulties
Lochman & Wells (2002)	US	Primary data via questionnaires	Contextual social-cognitive model	Changes in social-cognitive processes, schema and parenting practices, even among high-risk boys, can have a meaningful impact on later negative outcomes	183 boys, their parents, and teachers	Primary and middle schools	0%	NR - 4th and 5th grade	Boys were in the top 22% in teachers' ratings of children's aggressive and disruptive behaviours

<b>Study details</b>	<b>Country</b>	<b>Design</b>	<b>Model</b>	<b>Findings</b>	<b>Sample size</b>	<b>School level</b>	<b>% of females</b>	<b>Age M(SD)</b>	<b>Type of sample</b>
Lopez & DuBois (2005)	US	Primary data via questionnaires and secondary data via school records.	Integrative model of the effects of peer victimization and peer rejection	Peer victimization and peer rejection contribute independently to problems in emotional, behavioural, and academic adjustment	508 children	Middle school	48%	11 to 13-year olds	6 <sup>th</sup> and 7 <sup>th</sup> graders attending a Midwestern community school.
Lorber & Egeland (2011)	US	Primary data from longitudinal study via observations and questionnaires	Mutual exacerbation model	Mother's negative-infant difficulty pattern may support the development of significant early conduct problems	267 children, mothers and teachers	Neonatal days 7 and 10; at 3, 6, 24, and 42 months; kindergarten and 1 <sup>st</sup> grade	45.3%	From 7 days to 1 <sup>st</sup> grade Mothers 20.5(6.6)	High-risk urban sample, receiving prenatal care from a public health clinic
Lyons & O'Connor (2006)	Ireland	Primary data via questionnaires and interviews	Integrated model of the nature of challenging behaviour	With knowledge of context and our expectations a careful balancing and negotiation is needed to cope successfully with challenging behaviour	290 children 29 teachers	Primary	NR for questionnaires 53% for interviews	Aged 9 and 12.	Disadvantaged urban area.

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
McDermott et al (2001)	US	Primary data via questionnaires and observations with latent structure analysis	Structural model of student performance	Verbal and non-verbal learning play no appreciable role in behavioural outcomes whereas knowledge of motivational and disciplined behaviour does	1268 children and their teachers, parents/guardians	Primary through to secondary	50%	6 to 17 years of age	A nationally representative (stratified) sample
Mooij (1999)	Netherlands	A multilevel theorising article	A multilevel model	Social behaviour relates to the degree to which a child is a victim or perpetrator of aggression or vandalism inside and outside school as well as a perpetrator of criminal behaviour in later school years	N/A	All	N/A	N/A	All
Nelson et al (1999)	US	Primary data via questionnaires in a longitudinal study	Temperament theory	Negative emotionality is positively related to externalizing and internalizing problems and negatively related with positive social behaviours	75 children, their teachers and parents	Pre-school through to primary	52%	NR. Ranged in age from 7 years 1 months to 8 years 9 months	Suburban middle -class schools.
Nie & Lau (2009)	Singapore	Primary data via an on-line survey	Self-determination theory	Behavioural control is a significant negative predictor of classroom misbehaviour	3196 children	Secondary	51%	15.5 years	117 grade 9 classrooms in 39 schools.



Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Nolan (2011)	US	Primary data via ethnographic interviews and observations	Oppositional behaviour theories	Alienated students seek social and psychological benefits through oppositional behaviour, hence need positive education and/or political engagement	30 children 20 staff and 2 law enforcers	Secondary	N/A	Unclear	One school in a low-income neighbourhood in the Bronx, 99% Black and Latino, focussing teacher characterised 'chronic troublemakers'
Norwich & Rovoli (1993)	UK	Primary data via interviews, observations and questionnaires and secondary data via school records	Affective factors and learning behaviours	In a lesson a child's beliefs about what will make it hard to learn and their judgments about carrying-out certain learning behaviours are likely to influence their plans to engage in relevant learning behaviours	28 children	Secondary	43%	NR. 12 aged 11 to 12 years, 8 aged 12 to 13, and 8 aged 13 to 14	Two children of average and two of low maths attainment in each of eight classes were selected maths teachers. Those that had difficulties in reading were excluded
Núñez & León (2015)	Spain	A review / opinion paper	Psychological needs theory	Teachers who support autonomy improve student academic performance, are more creative and better adjusted, engage more in school, and feel less stress	N/A	All	N/A	N/A	All

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Olvera (2008)	US	A review of Emotional Disturbance (ED) literature and related models, and an evaluation of ED assessment questionnaires	Hypothesis Testing (HT) model and a range of anti-social behaviour theories	Anti-social behaviour is attributed to: lack of impulse control, parent-child relationship problems, temperament, being products of their environment, psychopathic mind, attachment problems DSM-IV-TR Adapted Screening Interview was thought to be helpful in screening childhood antisocial behaviour	9 school psychologists	Primary	N/A	N/A	Those children with emotional difficulties
Pingault et al (2015)	Canada	Primary data via interviews, questionnaires and observations	Childcare and social behaviour	Receiving childcare correlates to being less shy and socially withdrawn, more oppositional and aggressive. Differences dissipated with age	1,544 children their teachers and PMK	From pre-school through to end of primary	51.4% children PMK > 98%.	NR. Data from 5 months, 1½, 2½, 3½, 4 to 10 and 12 years	A birth cohort from Quebec, who had relevant child care and behavioural data
Powell & Tod (2004)	UK	Systematic review of learning behaviour literature	A framework reflecting the complexity of variables that influence learning behaviour	Theories that may contribute to understanding factors involved in learning behaviours are affective (self /engagement), cognitive (curriculum access) and social (social/ participation).	46 studies	All	NR	NR. Age range 3-16 years	All

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Richards et al (1984)	US	Primary data via interviews and questionnaires	Kohlberg's theory of moral development	Those that reason at higher moral levels and girls who have lower moral reasoning have less conduct problems	87 children and their teachers	Two primary and one secondary	52%	NR 4 <sup>th</sup> , 6 <sup>th</sup> and 8 <sup>th</sup> graders	Rural school children predominantly from lower class homes with as many stages of moral reasoning as possible.
Richards (1989)	US	Primary data via interviews and questionnaires	Two causal models of moral reasoning and conduct	Higher social class, being female, and scoring toward either extreme of the moral maturity continuum have less conduct problems.	60 children and their teachers	Two primary and one secondary	45%	NR. 4 <sup>th</sup> and 8 <sup>th</sup> graders	Rural school children predominantly from lower class homes with as many stages of moral reasoning as possible.
Slee (2014)	Australia	An opinion essay	Furlong's analysis of student disaffection and links to social theory and bio-politics.	We need to test our theorising across disciplinary networks to gauge the depth and quality of analysis in response to student disaffection	N/A	All	N/A	N/A	Targets children with behavioural and cognitive defectiveness.
Smith (2010)	US	Primary data via questionnaires and secondary data via school records	Social learning theory	Interventions to reduce or prevent suspensions must be undertaken in the school context	385 parents or guardians	Secondary	NR	NR. 10 <sup>th</sup> graders	Children with disciplinary referrals from schools with the largest student bodies in a northwest Louisiana school district

Study details	Country	Design	Model	Findings	Sample size	School level	% of females	Age M(SD)	Type of sample
Sullivan & Hirschfeld (2011)	US	Secondary data via James Comer's School Development Program intervention	Social Development Model (SDM) a life course theory	Prosocial and anti-social development is conditioned by the interaction between students and social contexts (schools, families, community programs)	2,014 children	Middle schools	53%	11.5(.68)	Socially and economically disadvantaged minority children.
Svendham (1994)	Sweden	Primary interviews, sociograms, network maps and questionnaires	Social network theory	Children's fragmented social networks influence behaviour problems in school	190 children their teachers and parents	Primary	44%	NR. M=11.9 at grade 5, M=12.9 at grade 6	Children with behaviour problems in school
Tremblay (2010)	UK	Opinion and review	Developmental trajectories of disruptive behaviours	Deficits in using socially accepted behaviours are intergenerational, based on complex genetic and environmental contributions. Prevention requires early, intensive and long-term support to parents and children	N/A	All	N/A	N/A	All

NR - Not reported; N/A – Not applicable; PMK - Person most knowledgeable about the child.

Note 1 Eliciting – parents engage in activities chosen by the child, talked to them about their feelings, and allows them to ask questions

## Review 1 synthesis and framework

There are a myriad of factors which can explain pupil behaviour in schools. The most extensive systematic review of the theories of behaviour in education contexts was conducted by Powell and Tod (2004) from Canterbury Christ Church University College which considered influences such as family, community, policy, and relationship with learning. Figure 4 which accompanies this review highlights the convergence of influences over behaviour. This diagram is designed to demonstrate the influence of life and educational events on individual pupils. If we consider the example of a pupil in Year 8 who has experienced the death of a close and influential relative. If prior to this the pupil was somewhere in the top right quadrant (optimal position), which indicates that the influences on behaviour are positive, the bereavement can start having a negative influence on behaviour and the pupil may gradually move through the quadrants (to bottom right then bottom left) and thus the influence on behaviour would be moving from positive to negative. These changes may be subtle but over time they have a substantial effect on the pupil's wellbeing or academic development. By using this model teachers may be more attuned to events which may have an influence over the pupils in their care. As we have mentioned earlier, by becoming aware of events before they become extreme then there is more chance of mitigating any negative change and thus having more chance of keeping the pupil in positive zones, both in terms of overall school behaviour and overall influences.

Loeber and Farrington (1999) state that intervention is 'never too early and never too late' when discussing 'serious and violent juvenile offenders'. However, in an education setting when behaviour is manifesting in extreme situations the preventative aspect is no longer a viable option at that point in time. The key aspect is to prevent the behaviour from reaching that situation, if possible. Being in an informed position where one can be aware of variables starting or continuing to affect a pupil's life situation is key to understanding and being effective in behaviour management.

# Behaviour & Influence

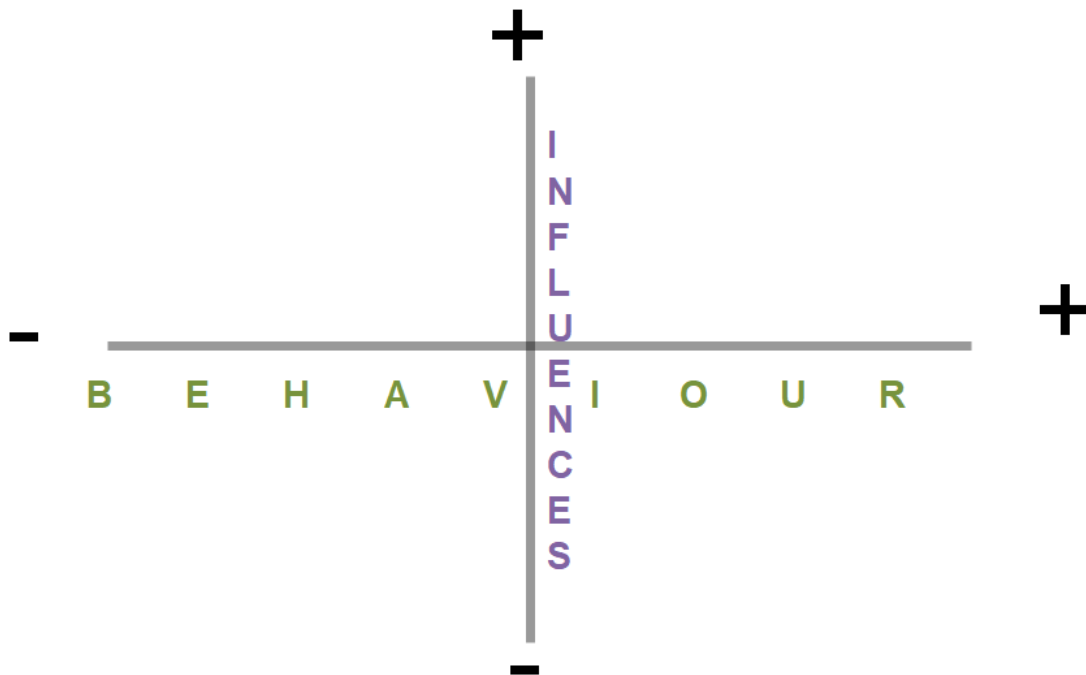


Figure 4. The link between positive and negative school behaviour and influences.

Most people will have experienced challenging situations in their life and resilience becomes important in facilitating their ability to handle adversity. To take again the example of the Year 8 child who had experienced a bereavement: their resilience is likely to also influence their behaviour, but also other things that may enable their resilience like a strong social support network and access to counselling might mean that such positive influences are able to mitigate against adverse events. It is important therefore to appreciate the multiplicity of influences on behaviour and how the same new negative influence will interact with other influences in different ways for different individuals. Being able to cope with adverse situations helps ensure that negative behaviour manifestations will be less common. In contrast, developing a negative self-image may be the result of different events that occurred during the early life cycle, that in turn may lead to or exacerbate negative behavioural manifestations in response to additional adversity. Thus, events outside school may impact children's behaviour directly but also via links to the perception of one's ability in school that can amplify the stress and distress experienced. There are influences on behaviour which teaching staff can affect directly, as well as others where there is a potential for teaching staff to influence or advise and finally a third category which indicates that the influences may be outside the purview of teaching staff. This is indicated in Table 2.

Pupils who are experiencing challenging situations at home and/or the community will then find that difficulties with learning or coping in the school environment may be more profound. Negative attributions for success and failure in learning (Chodkiewicz & Boyle, 2016, 2017) may become more entrenched which may then lead to a withdrawal from learning. In some cases this can lead to the void being filled with negative behaviour, which is sometimes linked to maintaining self-esteem and social standing with peers. This can become more pronounced in the later primary and secondary years where social skills and peer acceptance becomes more valuable (Bosnjak, Boyle & Chodkiewicz, 2017).

Table 2. Themes evidenced as affecting school behaviour in Review 1 included studies. Organised by level of school influence.

<b>Aspects to Manage Directly</b>	<b>Aspects to Identify and Influence</b>	<b>Aspects to be Aware of</b>
<b>Relationships (G)</b>	<b>Relationships out of school (G)</b>	<b>Home life (G)</b>
Teacher interest (G)	Choice of peers (G)	Sociometric Status (G)
Social Competency (G)	Relationship with Others (G)	Family Functioning (G)
Relationship with Peers (G)	Peer Group Perception of Consequences (M)	Witnessing Violence (G)
Relationship Development (G)	Family relationships (G)	Parental Mental Health and Wellbeing (G)
Teacher Connection (G)	Social Support (G)	Parental Education Level (G)
		Parental view of Education (G)
<b>Teaching and Learning (G)</b>	<b>Home life (G)</b>	Abuse (G)
Academic Achievement (G)	Discipline (G)	
Learning potential (G)	Parental Involvement in Homework (G)	<b>Culture (M)</b>
Educational opportunity (G)		Culture Conflict (M)
Connection to Curriculum (M)	<b>Stress and coping (G)</b>	Acculturation (M)
	Adverse Life Events (G)	
<b>School organisation (G)</b>	Adaptive Functioning(G)	
Exclusions/Suspensions (L)	Emotional Functioning (G)	
School Environment (E)	Behavioural Functioning (G)	
School Ethos/Policies (G)	Coping and resilience (G)	
School Transitions (L)		
School leadership (G)		
Behaviour management approach (G)		
Definition of appropriate behaviour (M)	<b>Behaviour out of school (M)</b>	
School & community relations (G)	Anti-Social Behaviour (G)	
Class Size (E)	Incarceration (L)	
<b>Attitudes and self-concept (G)</b>	<b>Ability (M)</b>	
Educational Motivation (G)	Intellectual Ability (M)	
Academic Expectations (G)		

Aspects to Manage Directly	Aspects to Identify and Influence	Aspects to be Aware of
Attitudes to Learning (G)	<b>Nutrition (E)</b>	
Locus of Control (G)	Breakfast (E)	
Labelling (G)		
Happiness (M)	<b>Out of school support (G)</b>	
Personal Perception of Consequences (G)	Social and welfare services (G)	
Self-evaluation (M)		
Emotional regulation (M)		
Attribution for attainment (G)		

Key to quality of research (not strength of association): E = Excellent; G = Good; M = Medium; L = Low

It follows that one of the main advantages teachers might hold when working with students is a knowledge of the pupil and his or her situation. Research suggests the positive impact on classroom behaviour when teachers know their students well (Sammons et al, 2016; Sizer, 1992), although, the implication is that information needs to be sought by teachers and willingly shared by pupils and parents. In settings where multiple adults frequently work with individual pupils, effective communication to colleagues by students' key adults is important. If we consider Figure 4 we can understand that positive influences affect behaviour in a strengthening manner. However, negative influences may lead to poor behaviour, and many of the factors in Table 2 may operate in both directions depending on their context. A good example is the quality of peer relationships or teacher-pupil relationships. Many people will move in and out of all four quadrants in Figure 4 over time, depending on life events and their ability to withstand adverse situations. A useful part of this review is the recognition that if we are able to understand that a child may be vulnerable and at risk of moving into a zone of difficulty, because of a particular life event (e.g. parental death, unable to grasp new education concept), it may be possible to intervene before a more chronic behaviour pattern emerges. More broadly, research suggests that responding to monitoring behaviour is important to limit long-term outcomes such as peer rejection and school failure if behavioural difficulties become more entrenched (Petersen et al., 2015; Reinke et al., 2009; Tyler-Merrick & Church, 2012).

## Review 2 and 3 Methods

### Review 2 aims

Review 2 focused on the effectiveness of approaches to classroom behaviour management. Specific research questions include:

- What types of classroom management approaches are most effective in improving attainment, learning or behaviour of:



- all pupils?
- pupils who exhibit challenging behaviours?
- Which components of approaches, or theories that explain their mechanisms, are most promising?

### Review 2 Inclusion criteria

*Population:* School-aged children 4-18 years of age. Populations that were exclusively in post-16 education institutions (e.g. Further Education Colleges) or preschool were excluded. The intervention could be delivered via school staff as well as directly to children. Outcomes needed to relate to schoolchildren in the setting. The educational setting could include mainstream and specialist settings. Child participants were either all students in the setting or a specific group (e.g. conduct disorder, or children with SEN).

*Intervention:* An approach, strategy or programme that primarily targets improving student behaviour in the classroom and the approach either is delivered in the classroom or trains staff who then take action in the classroom, rather than across the school setting.

*Study design:* Any randomised controlled trial or quasi-experimental primary research. Quasi-experimental research relates to comparisons of intervention and control group when the allocation to group is not randomised. Studies with single intervention groups and pre-post-intervention measures (no control group) were excluded. Systematic reviews that may have relevant studies were retained and their included studies located.

*Outcome:* Included studies must have had at least one outcome that would fit the following categories: behaviour, attainment, other learning outcomes.

*Date:* Any.

*Language:* English only.

*Publication type:* Peer-reviewed research. Dissertations were excluded.

### Review 3 aims

Review 3 focused on the effectiveness of school-wide approaches to behaviour management. It included both popular whole school approaches to behaviour management, as well as effective leadership practices that promote a school culture of positive behaviour.

Review 3 aimed to answer the following research questions:

- What types of whole school approaches to behaviour management are most effective in improving learning outcomes and behaviours of:
  - all pupils?
  - pupils who exhibit challenging behaviours?

- Is there evidence that whole school approaches (review 3) are more or less effective than classroom-based approaches (review 2)?
- What gaps for randomised controlled trials and quasi-experimental studies exist for frequently used whole school approaches to behaviour management?

### Review 3 inclusion criteria

Articles were selected according to the following inclusion criteria:

*Population:* School-aged children 4-18 years of age. Populations that were exclusively in post-16 education were excluded. The intervention may have been delivered to school staff to then apply as a whole school approach rather than children receiving the intervention but outcomes needed to relate to schoolchildren in the setting. Special Schools and schools with particular intakes (e.g. selective on ability, vocational curriculum) were included, therefore whole school approaches may be focused on specific types of children, e.g. those with social, emotional and mental health difficulties.

*Intervention:* An approach, strategy, programme or policy that primarily targets behaviour at the whole school level. Whole school approach meant that either study authors identified it as a whole school approach, every classroom in the school applied the same intervention or elements incorporated behaviour beyond the classroom, e.g. school grounds or behaviour before school. The approach was either delivered in the school or trained at least some staff with the intention that they effect improvement at a whole school level. The interventions may span both staff action and leadership practice. A whole school approach is very likely to specify particular action in the classroom. While a single study may fit the inclusion criteria for both Review 2 and Review 3, it would only be included in both if it focused on the impact of the classroom elements as well as the remainder of the whole school approach.

*Study design:* Any randomised controlled trial or quasi-experimental primary research. Studies with single intervention groups and pre- post-intervention measures (no control group) were excluded. Systematic reviews that may have relevant studies were retained and their included studies located.

*Outcome:* Included studies must have at least one outcome that would fit the following categories: behaviour, attainment, other learning outcomes.

*Date:* Any

*Language:* English only.

*Publication type:* Peer-reviewed research. Dissertations were excluded.

### Search strategy for both Reviews 2 and 3

Because of the similarities between the inclusion criteria for Review 2 and 3, a single search, study selection, data extraction and synthesis process was followed for both reviews. The database searches were designed and run by an information specialist. The search combined terms for study design, behaviour, school, schoolchildren and intervention.

Between 28th September and 2nd October 2018 we searched the databases ERIC, Education Research Complete and the British Education Index (via EBSCOhost), the Australian Education Index and ASSIA (via ProQuest), PsycINFO (via OvidSp), Social Science Citation Index (via Web of Science), the Cochrane Library and the Campbell Library.

An example search strategy can be seen in appendix 3.

Forwards and backwards citation chasing was conducted for all articles included from the database search. The included studies from relevant systematic reviews identified from the search were also screened for any relevant additional studies.

We also searched the following websites for potentially unpublished literature: Education Endowment Foundation, What Works Clearing house, Department for Education, The Schools, Students and Teachers Network, Devon LEA, Parentkind.

Key journals checked were those where multiple included studies were published.

### Study Selection for Reviews 2 and 3

References located by the search were uploaded to reference management software (Endnote X8) and duplicate studies were removed. Relevant studies were identified in two stages based on the inclusion/exclusion criteria given above. First, independent double screening of titles and abstracts for each record was conducted by two reviewers. Full texts of records that appeared to meet the inclusion criteria on the basis of titles and abstracts were then obtained wherever possible via the University of Exeter online library or web searching. Each full text article was screened independently by two reviewers (SBC and RK). Reasons for exclusion at this stage were recorded. Disagreements were resolved by a third reviewer as necessary (DM). During full-text screening any relevant systematic reviews were excluded after all relevant included studies from these reviews were screened.

### Data Extraction for Reviews 2 and 3

Data was extracted from included studies in table format to aid synthesis and comparison across included studies. A data extraction form was developed and piloted. Data on article details, sample, intervention, outcome measures, findings and study quality were extracted

into Microsoft Office Excel 2013 (Microsoft, Redmond, CA, USA) by either RK, SBC or LG and checked by DM.

### Critical appraisal

We appraised studies using the Joanna Briggs Institute checklists for randomised and quasi-experimental studies (Joanna Briggs Institute, 2017a, 2017b). After piloting the Joanna Briggs Institute checklists and the RoB 2.0 (Higgins et al., 2016) and ROBINS-I (Sterne et al., 2016) checklists, the former was selected because of ease of use and relevance to included studies. The quality appraisal of studies were used to evaluate risk of bias and study quality and were not used to exclude papers.

### Categorisation of interventions and outcomes

During data extraction, interventions were categorised according to similarities in terms of broad intervention type and intervention content. The labels and definitions of these intervention categories were developed using the descriptions of interventions in the included studies. The categorisation was primarily used to organise the quantitative synthesis and ensure that groups of studies considered together included interventions that were sufficiently similar.

Given that included studies needed to have at least one outcome measure that fitted either behaviour, academic outcomes or other learning outcomes, outcomes were categorised using these three categories. We soon found that because interventions focused on school or classroom behaviour, there were rarely any outcomes that did not fit these categories. When this occurred they were categorised as “other”. Social relationships were an exception, appearing as an outcome in several studies.

### Synthesis

We categorised included studies by: universal or targeted approaches; type of classroom management or whole school approach; and school level. We applied these categories to synthesise groups of similar studies (e.g. secondary school, teacher training interventions, targeting all pupils) describing the studies' findings in relation to behaviour, attainment, other learning and any other outcomes beyond these primary review outcomes. Wherever possible we calculated effect sizes for behaviour outcomes for each study to compare effects in a consistent manner.

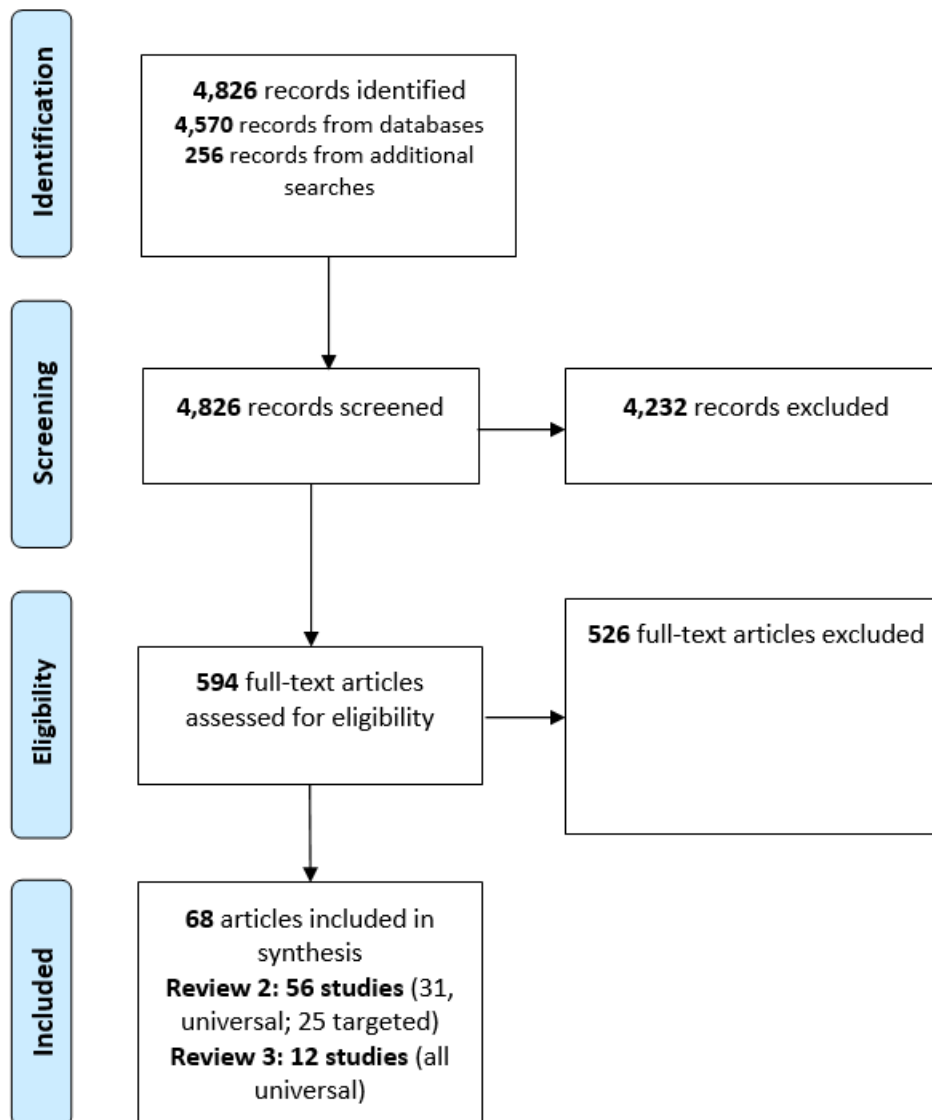
Differences between intervention and control group means reported at the first time-point after the intervention was completed were analysed. Cohen's effect size (d), the standardised mean difference, was reported for each outcome measure category. The effect sizes and 95% confidence intervals were calculated using the mean, standard deviation and

the sample size for the intervention and control groups or, if any were not reported, statistics that could be used to derive these (e.g. confidence intervals). When two or more measures that assessed the same outcome category were reported in a study, the effects were combined into one composite effect for that outcome by calculating the effect size for each measure and using the mean; we calculated the standard error for this composite effect in the usual way using smallest sample size across measures. Scores were analysed so that a positive effect size indicates some degree of improvement and a negative effect size indicates some degree of deterioration in order to avoid confusion between outcomes that might *increase* prosocial behaviour or *reduce* misbehaviour. Where effect sizes and 95% confidence intervals could not be calculated by reviewers, brief narrative findings are reported with reference to the included article.

### **Review 2 and 3 Findings**

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram in Figure 5 summarises the process of study selection. Initial database searching gave 4570 records to title and abstract screen. Four hundred and eighty four of these records were considered relevant and their full texts were retrieved and screened, leading to 50 included studies. Additional searches led to 110 more full text records screened and 18 more studies included. This gave a total of 68 included studies, 56 of which fit Review 2 as classroom-based interventions; the remaining 12 fit Review 3 as whole-school interventions.

Figure 5. PRISMA flow diagram showing study selection process for Reviews 2 and 3



### Review 2 descriptive synthesis

Synthesis for Review 2 is split at the broadest level between those studies that are delivered to universal student populations and those that are delivered to targeted student populations.

### Universal interventions

The summary details of the included studies (n=31) can be seen in Table 3. The included studies were published from 1974 to 2018, and most were carried out in the USA (n=19). Other countries included UK, Ireland, Norway, Belgium, Canada, Germany, Hong Kong and the Netherlands. The study designs included RCTs (n=24) with one (n=20) or two (n=4) treatment groups, and quasi-experimental (n=7) studies with one (n=6) or two (n=1)

treatment groups. The control groups were mostly treatment as usual (n=14) and waitlist (n=10). Other controls usually involved offering similar levels of attention to participants without exposing them to the actual intervention. Most studies used cluster randomisation (n=28). Children were participants in all but one study, with n=20 including both children and teacher participants. In addition, one study included wider school staff as participants. Sample sizes of participants ranged from 41 to 8350 children, and 6 to 469 teachers. Included schools were mostly elementary or primary levels (n=28). The remaining studies included samples from middle schools, high schools or secondary schools (n=3). Percentage of female children ranged from 24% to 73%, while the percentage of female teachers ranged from 60% to 100%. Where reported, children's mean age ranged from 4.2 to 13.5 years.

Table 4 gives details of the interventions used in the included studies. The main categories of intervention were teacher training including reward systems (e.g. IYTCM) which mainly focussed on classroom management skills; teacher training that did not include reward systems (e.g. coaching teachers); reward systems (e.g. GBG), all of which involved training teachers. Teacher training interventions were primarily focused on upskilling teachers in their behaviour management skills and practice, whereas reward system interventions included training on how to use the system; that is they focused on implementing a particular reward system rather than explicitly on the teacher developing skills and then applying them as they wished. Fidelity was assessed in 19 of the included studies.

Quality appraisal ratings for each included study are shown in Tables 5 and 6, for RCTs and quasi experimental studies respectively. RCTs typically used true randomisation, analysed participants in the groups they were allocated to (i.e. intention to treat), measured outcomes reliably and used appropriate trial design and analysis. Perhaps unsurprisingly RCTs less often blinded participants, assessors and those delivering interventions to treatment assignment. Although reporting was often unclear, it was rare for studies to be at risk of bias across a wide range of criteria. Only Rogeness (1977) scored particularly poorly, perhaps because of the age of the study, although the findings from this study are not particularly strong regardless. Quasi-experimental studies scored well (every quality criteria either evident or unclear), notwithstanding their non-randomised allocation of groups.

Table 3. Study characteristics for Review 2 studies for universal student samples

<b>Study</b>	<b>Country</b>	<b>Design</b>	<b>Type of control</b>	<b>Sample</b>	<b>Sample size</b>	<b>School level</b>
Aasheim 2018	Norway	Quasi	Waitlist	Teachers & students	1518 students.	Primary
Bartholomew et al (2018)	USA	RCT	Traditional sedentary academic lessons.	Children only	2716 children	Elementary
Caldarella 2018	USA	RCT	TAU (typical classroom management strategies)	Children & teachers	350 children	Elementary
Capella 2012	USA	RCT	Initial training, access to website	Mental health staff, teachers, children	12 MH staff. 36 teachers. 364 children.	Elementary
Dolan 1993	USA	RCT	TAU	First-grade classrooms, all students	864	Elementary
Humphrey 2018	UK	RCT	TAU	Children and teachers	3084 children	Primary
Evertson (1989)	USA	Quasi	TAU	teachers	29 teachers	Elementary
Evertson 1995	USA	Quasi	Waitlist	Teachers & students	46 teachers. 420 students in study 7. 423 students in study 8.	Primary (grades 1-6)
Fernandez (2015)	USA	Quasi	TAU	Teachers and their students	11 teachers, 118 students	Kindergarten and first grade classrooms in US public school
Ford 2018	UK	RCT	TAU	Children only	2188 children	Primary
Fossum 2017	Norway	Quasi	TAU/Waitlist ( <i>offered training 1 year later</i> )	Children	1218 children	Kindergarten
Gregory (2014)	USA	RCT	TAU	Teachers but their students were observed	1669 students	Middle and high schools



Han 2005	USA	RCT	No treatment	Children	166	Pre-Kindergarten
Hickey 2017	Ireland	RCT	Waitlist	Teachers & students	445 children.	Primary
Homer 2016	Hong Kong, China	RCT	Other - non-digital conventional school token point system	Children	120 randomised	Elementary
Hutchings 2013	UK	RCT	Waitlist	Teachers & students	12 teachers. 107 children.	Primary
Kamps 2015	USA	RCT	Waitlist	Children & teachers	159 teachers. 17 schools, average 382 students	Elementary
Leflot 2010	Belgium	RCT	TAU	Teachers & students	15 schools. 570 children.	Elementary
McGilloway 2010	Ireland	RCT	Waitlist	Teachers & children	11 schools, 22 teachers, 234 children at baseline	Infant
Murray 2018	USA	RCT	Waitlist	Teachers & students	Teachers: 97 baseline, Students: 1192 analysed.	Elementary
Okonofua 2016	USA	RCT	<i>The control exercise</i>	Teachers & students	2069 students recruited	Middle schools
Piowar 2013	Germany	Quasi	Module 1 only (presumed to represent the traditional instruction approach)	Teachers & children	37 teachers. 666 children.	Secondary
Reinke 2012	USA	RCT	Waitlist	Teachers & children	Teachers: 105 randomised. Children: 1817 randomised.	Elementary
Reinke 2018	USA	RCT	Waitlist business as usual control group	Teachers and students	105 teachers and 1817 students.	Elementary
Rogeness 1977	USA	RCT	TAU	Children	234	Elementary
Spilt 2013	Netherlands	RCT	Not reported	Children only	759 children from 47 classes	Elementary

Spilt 2016	Belgium	RCT	Not reported	Children & teachers	30 teachers. 570 children.	Elementary
Thompson 1974	USA	Quasi	TAU (no training/assistance)	Teachers & students	22 teachers. Students not reported.	Elementary
Wills 2016	USA	RCT	Control.	students and teachers	313 children and 169 teachers	Elementary
Wills 2018	USA	RCT	TAU, offered training in the spring to use the intervention	Children & teachers	Teachers: 193 recruited, Class size 18-25 students.	Elementary
Wills 2018	USA	RCT	Control.	students and teachers	157 teachers 324 children for details	Elementary

Table 4. Intervention details for Review 2 studies for universal student samples

Study	Intervention name	Intervention category	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Aasheim 2018	IYTCM	Teacher training (including reward system)	All 1st-3rd grade teachers & after school service staff	<i>Experienced &amp; qualified group leaders</i> (see page 7 paragraph 2 for further details)	School year (8-9 months between pre & post assessment)	No
Bartholomew 2018	I-CAN!	Physical Activity	Children	Teacher	Monday to Friday	Not mentioned
Caldarella 2018	CW-FIT	Teacher training (including reward system)	Teachers receive training and given scripted lessons to teach children.	Research staff train teachers, teachers deliver to children.	1 academic year	Yes - fidelity
Capella 2012	BRIDGE	Teacher training (no reward system)	Consultants, teachers	Unsure who delivers to consultants. Consultants deliver to teachers. Teachers deliver to children.	1 year	Mentioned in abstract but not reported in detail
Dolan 1993	GBG Good Behaviour Game (GBG)	Reward system	Teachers given training to implement intervention to children	Teacher delivers to class. Unclear who delivers training for teachers.		No

Study	Intervention name	Intervention category	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Evertson (1989)	A School-Based Training Programme	Teacher training (including reward system)	Teachers	Trainers	All teachers in this experiment were observed on six occasions, four times after the first work-shop and twice after the second workshop in mid-October.	No
Evertson 1995, COMP	Classroom Organisation and Management Program (COMP)	Teacher training (including reward system)	Teachers	Certified COMP workshop leader	Tests taken in fall and spring terms. 7-8 month interval between mathematics assessment.	No
Fernandez (2015)	Teacher-Child Interaction Training (TCIT)	Teacher training (including reward system)	Only the teachers	Both TCIT trainers were clinical psychologists, one a Parent-Child PCIT Master Trainer and the other a PCIT Level I Trainer	Coaching sessions continued until a teacher reached the pre-set skills mastery for both phases of the training or until 24 weeks of training had occurred based on resource limitations.	No
Ford 2018, Incredible Years Teacher Classroom Management programme	TCM	Teacher training (including reward system)	Teachers	Qualified group leaders	6 months	Yes
Fossum 2017	Incredible Years Teacher Classroom Management (IY TCM)	Teacher training (including reward system)	All staff, all children	<i>Experienced and qualified group leaders</i>	9 months i.e. whole school year	No

Study	Intervention name	Intervention category	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Gregory (2014)	My Teaching Partner-Secondary program (MTP-S)	Teacher training (no reward system)	Teachers - see intervention description.	Two coaches, who were both experts in teaching and adolescent development coached the teachers	academic year	Yes
Han 2005	Reaching Educators, Children, and Parents (RECAP)	Teacher training (including reward system)	Teachers, children, parents	Program consultants delivered teacher training & parent groups - two masters-level clinicians	Pre-treatment assessments September/October, post-treatment assessments April/May	Yes - fidelity
Hickey 2017	IYTP	Teacher training (including reward system)	Teachers	Trained facilitators	5 months	Yes - facilitator-reported fidelity. Adherence not independently validated.
Homer 2016	ClassDojo	Reward system	Children	Teacher	16 weeks	No
Humphrey (2018)	Good Behaviour Game (GBG)	Reward system	Teachers receive training, then implement in classroom	GBG coaches	2 years	Yes
Hutchings 2013	IYTCM	Teacher training (including reward system)	Teachers	<i>Two trained leaders, a certified program mentor (the first author) and a trained program leader</i>	5 months	Yes - fidelity
Kamps 2015	Class-Wide Function-Related Intervention Teams (CW-FIT)	Reward system	Children	Project staff deliver training. Teachers deliver to children.	5 months (October to March)	Yes - fidelity
Leflot 2010	GBG	Reward system	Children	Teacher	7 months	Yes - "implementation"

Study	Intervention name	Intervention category	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
McGilloway 2010	IYTCM	Teacher training (including reward system)	Teachers	Not reported	6 months	No
Murray 2018	IYTCM	Teacher training (including reward system)	Teachers	Trained group leaders	6 months	Yes = fidelity
Okonofua 2016, Empathic Discipline	Empathic-Mindset Intervention	Teacher training (no reward system)	Teachers	Online module	First module midway through fall term. Second module 2 months later.	No
Piowar 2013	Kompetenzen des Klassenmanagements (KODEK) "classroom management competencies"	Teacher training (including reward system)	Teachers	Unclear	Unclear - possibly 1 academic year	Unclear - teacher knowledge on classroom management, competencies in classroom management (p 5&6)
Reinke 2012 IY-TCM	Incredible Years Teacher Classroom Management (IY TCM)	Teacher training (including reward system)	Teachers	2 doctoral-level IY TCM group leaders who were supervised by the program developer; one of these trainers also served as a coach.	1 academic year	Yes - fidelity

Study	Intervention name	Intervention category	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Reinke 2018	IYTCM	Teacher training (including reward system)	Teachers trained	Two doctoral-level IY TCM group leaders who were supervised by the program developer conducted the full-day trainings; one of these trainers also served as a coach.	Approx 6 months	Yes
Rogeness 1977 counselling	Behaviour modification & counselling	Reward system	Initially teachers, then children	Mental health consultants	1 academic year, replicated for 2 years	No
Spilt 2013	Good Behaviour Game (GBG)	Reward system	Children	Teachers	2 years	No
Spilt 2016	Good Behaviour Game (GBG)	Reward system	Children	Teachers	1 academic year	Yes - implementation
Thompson 1974	Training in behaviour management	Teacher training (including reward system)	Teachers	Authors/Researchers	Unclear	No
Wills 2016	Class-Wide Function-related Intervention Teams (CW-FIT)	Reward system	Teachers trained to deliver intervention to selected children	Teacher delivered after training by coaches	About 7 months.	Fidelity
Wills 2018	Class-Wide Function-related Intervention Teams (CW-FIT)	Reward system	Teachers trained to deliver intervention to selected children	Teacher delivered after training by coaches	About 7 months.	Yes
Wills 2018	Class-Wide Function-Related Intervention Teams (CW-FIT)	Reward system	Teachers & children	Teachers	October to March. Replicated over 4 years.	Yes - fidelity





Table 5. Review 2 Universal Interventions RCT quality

Study/Criteria	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
Bartholomew (2018)	U	U	Y	N	N	N	Y	U	Y	Y	U	Y	Y
Caldarella (2018)	Y	N	U	U	N	N Note 1	N	U	Y	Y	U Note 1	Y	Y
Capella (2012)	Y	N	Y	U	N	N Note 1	N	Y	Y	Y	U Note 2	Y	Y
Dolan (1993)	Y	U	Y	N	N	N	Y	Y	Y	Y	N	Y	Y
Ford (2018)	Y	Y	Y	N	N	Y	Y	N	Y	Y	Y	Y	Y
Gregory (2014)	N	N	Y	N	N	Y	Y	Y	Y	Y	Y	U	Y Note 4
Han (2005)	U	U	Y	U	N	N Note 3	U	Y	Y	Y	N	Y	Y
Hickey (2017)	Y	Y	Y	N	N	Y Note 5	U	N	Y	Y	Y	Y	Y
Homer (2016)	N	U	U	N	N	N	Y	Y	Y	Y	N	Y	Y
Humphrey (2018)	Y	U	N	N	N	N	U	N	Y	Y	Y	Y	Y
Hutchings (2013)	Y	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y
Kamps (2015)	Y	Y	U	U	N	N	U	Y	Y	Y	Y	U	Y
Leflot (2010)	Y	U	Y	N	N	N	U	N	Y	Y	Y	Y	Y
McGilloway (2010)	Y	Y	U	N	N	Y	U	N	Y	Y	Y	Y	Y
Murray (2010)	Y	U	U	N	N	Y	N	U	Y	Y	Y	Y	Y
Okonofua (2016)	U	U	U	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Reinke (2012)	Y	U	Y	N	N	Y	Y	U	Y	Y	Y	Y	Y
Reinke (2018)	U	N	Y	N	N	U	U	Y	Y	Y	Y	U	Y
Rogeness (1977)	N	N	N	U	N	N	N	N	N	Y	N	U	U
Spilt (2013)	Y	U	Y	N	N	N	Y	N	Y	Y	N	Y	Y
Spilt (2016)	Y	U	N	N	N	N	Y	N	Y	Y	Y	Y	Y
Wills (2018)	Y	Y	U	N	N	N	Y	N	Y	Y	Y	U	Y
Ward (2013)	Y	U	Y	N	N	N	U	U	Y	Y	Y	Y	Y
Wills (2016)	Y	Y	Y Note 6	N	N	N	U	Y	Y	Y	Y	Y	Y
Wills (2018)	Y	N	N	N	N	N	U	Y	Y	Y	Y	Y	Y

Y= Yes, N=No, U=Unclear.

1. Was true randomization used for assignment of participants to treatment groups?  
 2. Was allocation to treatment groups concealed?  
 3. Were treatment groups similar at the baseline?  
 4. Were participants blind to treatment assignment?  
 5. Were those delivering treatment blind to treatment assignment?  
 6. Were outcomes assessors blind to treatment assignment?  
 7. Were treatment groups treated identically other than the intervention of interest?  
 8. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?  
 9. Were participants analysed in the groups to which they were randomized?  
 10. Were outcomes measured in the same way for treatment groups?  
 11. Were outcomes measured in a reliable way?  
 12. Was appropriate statistical analysis used?  
 13. Was the trial design appropriate, and any deviations from the standard RCT design (individual randomization, parallel groups) accounted for in the conduct and analysis of the trial?

Note 1 Teachers - N. Observers - U.

Note 2. Teachers - N. Observers - Y.

Note 3. N - teacher report. U - parent report.

Note 4. Randomised to district level only

Note 5. All except one researcher

Note 6. Apart from significant differences at pre-test on teacher praise

Table 6. Review 2 Universal Interventions Quasi-experimental quality table

Study	1. Is it clear in the study what is the 'cause' and what is the 'effect' (i.e. there is no confusion about which variable comes first)?	2. Were the participants included in any comparisons similar?	3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	4. Was there a control group?	5. Were there multiple measurements of the outcome both pre and post the intervention/exposure?	6. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?	7. Were the outcomes of participants included in any comparisons measured in the same way?	8. Were outcomes measured in a reliable way?	9. Was appropriate statistical analysis used?
Aasheim (2018)	Y	Y Note 1	U Note 2	Y	Y	U	Y	Y	Y
Evertson (1989)	Y	Y	Y	Y	Y	Y	Y	Y	Y
Evertson (1995)	Y	U	U	Y	Y	U	Y	Y	Y
Fernandez (2015)	Y	Y	Y	Y	Y	Y	Y	U Note 3	U
Fossum (2017)	Y	Y	U	Y	Y	Y	Y	Y	Y
Piwowar (2013)	Y	U	U	Y	Y	U	Y	Y	Y
Thompson (1974)	Y	U	U Note 2	Y	Y	U	Y	U	Y

Y= Yes, N=No, U=Unclear.

Note 1 One significant difference in ethnicity of students

Note 2 Educated in different schools

Note 3 No, regarding observational data on student behaviour

### Targeted interventions

The summary details of Review 2 studies aiming to improve the behaviour of targeted samples of students are shown in Table 7. Fifteen of the 25 studies are RCTs. The majority of comparator groups were treatment as usual. The targeted populations were typically children identified at risk of behaviour difficulties including disruptive behaviour, externalising problems and aggressive behaviour. Some samples comprised pupils at risk of or who struggled with emotional or behavioural disorders. Most samples were rated as at risk by teachers, rather than screened as above a threshold on a reliable and valid scale. Relatively few studies included samples with diagnosed behavioural concerns or school placement that indicated this level of behaviour problem. Sample sizes were often fewer than 100 participants. As for the universal studies included in Review 2, there was a predominance of elementary and primary schools as the school level, rather than secondary school. Indeed, no sample was exclusively in secondary or high school.

Table 8 shows details about the interventions for targeted students. As for the universal interventions, many interventions involved teacher training. Some focused on reward systems; these are programmes that involve the presentation of something such as a reward or praise, known as positive reinforcement, where the goal is to achieve a desired response, such as listening or responding to a request. Unique to the targeted interventions were those that focused on training student skills (e.g. functional behavioural assessments). Those who delivered the intervention were often trained and fidelity was often assessed.

Tables 9 and 10 show quality appraisal for these studies. The 15 RCTs rarely blinded treatment, delivery or assessment. Allocation to treatment groups was rarely concealed, or not reported clearly that it was concealed. Differences in the proportion of participants completing post-tests was less often described in these RCTs compared to universal interventions. Outcomes, data analysis and trial design were always free from bias. Quasi-experimental studies typically scored well on the criteria, free from bias on outcomes, control groups, analysis. Some studies did not have completed groups at post-test.

Table 7. Study characteristics for Review 2 studies for targeted student samples

<b>Study</b>	<b>Country</b>	<b>Design</b>	<b>Type of control</b>	<b>Targeted sample</b>	<b>Inclusion Criteria</b>	<b>Sample size</b>	<b>School level</b>
Axberg 2006	Sweden	Quasi	TAU	Children who displayed externalising behaviour problems (at risk)	Teachers reported child misbehaving to pupil's welfare conference	50 children (34 intervention 16 control)	Primary school
Benner 2012	USA	RCT	TAU	Children exhibiting externalising behaviour problems (at risk)	According to threshold on the Systematic Screening for Behaviour Disorders	70 children(n= 44 treatment, 26 control)	Elementary (Grades K-3)
Bishop 1996	USA	Quasi	TAU	Emotionally/behaviourally disturbed students (behavioural concerns)	Pupils in classes for emotionally/behaviourally disturbed.	48 (28 intervention 20 control)	High school and Elementary school
Breeman 2016	Netherlands	RCT	TAU	Children with psychiatric disorders (behavioural concerns)	Attending special primary education for children with psychiatric disorders	389 children (212 intervention 177 control)	Primary school (special education)
Cook 2017	USA	Quasi	TAU	Disruptive and off task classes (at risk)	Classes where observations revealed off task behaviour >30%	159 students	Elementary and Middle
Cook 2018a	USA	RCT	Attention control	All children in classes with high rates of disruptive behaviour/low attainment (a risk)	Classes where observations revealed <65% AET score	203 children	Middle school
Cook 2018b	USA	RCT	Attention control	All children in classes with high levels of disruptive/off-task behaviour (at risk)	Classes where observations revealed off task behaviour >20%	220 children	Elementary
Durlak 1980	USA	Quasi	Waitlist	Children with school adjustment problems (at risk)	Selected based on high scores on both AML and TRF	119 children (51 school A, 42 school B, 26 control)	Elementary
Eisenhower (2016)	USA	Quasi	Delayed control	children with behaviour problems (at risk)	Parent report or elevated scores for behaviour or internalising problems	97 children	Kindergarten

<b>Study</b>	<b>Country</b>	<b>Design</b>	<b>Type of control</b>	<b>Targeted sample</b>	<b>Inclusion Criteria</b>	<b>Sample size</b>	<b>School level</b>
Forman 1980	USA	RCT	Placebo and response cost technique	Children with aggressive behaviours (at risk)	Referral to school psychologist for aggressive behaviour	18 (14 intervention (7 cognitive, 7 response cost), 4 control)	Elementary
Forster 2012	Sweden	RCT	Other (Chemical Base Resolution Lies In Education (CHARLIE))	Students with externalising behaviour (at risk)	Teacher and SDQ ratings for externalising behaviour	100 (60 intervention 40 control) 38 schools (26 intervention 12 control)	Elementary
Fuchs 1990	USA	Quasi	TAU	Difficult-to-teach students (at risk)	Teachers one most difficult to teach student	43 students (10, 10, 11 intervention) (12 control)	Middle school
Gettinger & Stoiber (2006)	USA	RCT	Treatment as Usual	children whose behaviours were disruptive, harmful to others, and/or interfered with their learning (at risk)	Two children per class who exhibited challenging behaviours	70 children & 41 teachers	pre-kindergarten, kindergarten, and first-grade classrooms
Hops 1978	USA	Quasi	TAU	Children with disruptive, acting-out behaviour (at risk)	Not reported	54 children (27 intervention 27 control)	Elementary
Iovannone 2009	USA	RCT	TAU	Students with behavioural problems (at risk)	5= critical events on SSBD, persistent behaviour problems	245 children	Elementary, Middle, alternative schools
Kirkhaug 2016	Norway	Quasi	Waitlist	Children with clinical level externalising problems (behavioural concerns)	90 <sup>th</sup> percentile + on SESBI-R	83 (45 intervention & 38 control)	Elementary
Long 2018	USA	Quasi	TAU	Alternative school for students two or more grades behind and significant disciplinary problems (behavioural concerns)	Attending this school	73 students	Elementary
Palcic 2009	USA	RCT	TAU	ADHD (behavioural concerns)	Diagnosis and 98 <sup>th</sup> percentile + on Conners screening	43 students	Elementary

<b>Study</b>	<b>Country</b>	<b>Design</b>	<b>Type of control</b>	<b>Targeted sample</b>	<b>Inclusion Criteria</b>	<b>Sample size</b>	<b>School level</b>
Randolph & Hardage 1973	USA	RCT	No treatment	At risk of school dropout (at risk)	The six children screened by each teacher as having the highest drop- out potential in their class	90 children (30 in each intervention, 30 control)	Elementary
Reinke 2014	USA	RCT	TAU	Students with disruptive behaviour (at risk)	Top 15% of TOCA-C disruptive behaviour subscale	46 (23 intervention 23 control)	Elementary
Simonsen 2011	USA	RCT	TAU	Children exhibiting frequent behavioural problems (at risk)	Nominated by teacher as disruptive or 2+ office referrals in previous month	42 children (27 intervention 15 control)	Middle school
Stoiber 2011	USA	RCT	TAU	Children with challenging behaviours (at risk)	Two children per class teachers considered to have challenging behaviour	90 children 957 intervention 33 control)	Pre-k to 1st Grade
Trovato 1992	Canada	RCT	TAU	Children exhibiting behavioural concerns (at risk)	Identified by school evaluation as exhibiting behavioural concerns	77 children (58 children results analysed)	Elementary
van den Berg 2018	Netherlands	RCT	TAU	Externalising behaviour (at risk)	teacher-identified children with elevated levels of externalising behaviour at school	1569 children	Elementary
Weinrott 1979	USA	RCT	Placebo	Children described as 'acting-out' (at risk)	Teacher rated 2+ disruptive behaviours and 3+ distractible behaviours	20 teachers/student pairs (10 intervention, 10 control)	Elementary

Table 8. Intervention details for Review 2 studies for targeted student samples

<b>Study</b>	<b>Intervention name</b>	<b>Intervention category</b>	<b>Who receives the intervention?</b>	<b>Who delivers?</b>	<b>Duration</b>	<b>Fidelity assessed?</b>
Axberg 2006	Marte Meo Model	Teacher training	Teachers	Trained professional	Academic year	Unclear
Benner 2012	Behaviour Intervention	Teacher training	Principals, teachers, and staff	Trained professional	5 months	Yes (coaches/observers)
Bishop 1996	BTS	Student intervention	Students	Unclear	42 hours	Yes
Breeman 2016	GBG	Reward system	Teachers & students	Consultants	84	Yes
Cook 2017	(5):(1)	Teacher training	Students	Teacher	School year	Yes (by teacher carrying out intervention)
Cook 2018a	PGD	Teacher training	Students	Trained professional	School year	Yes (by teacher carrying out intervention)
Cook 2018b	EMR	Teacher training	Students	Trained professional	10-12 weeks	Yes (by teacher carrying out intervention)
Durlak 1980	Behavioural treatment and relationship treatment	Teacher training	Teaching Aides	Trained professional	2 months	Unclear
Eisenhower (2016)	Starting Strong	Teacher (and parent) training	Teachers and parents who then adapt their approach to the children and each other.	Two leaders facilitate both parent and teacher groups at each school	4 months	Yes
Forman 1980	cognitive restructuring and response cost	Student intervention	Students	Trained professional	Unclear (6 days?)	Yes
Forster 2012	COMET	Student intervention	Teachers	Trained professionals e.g clinical psychologists	5 months	Yes
Fuchs 1990	BC	Teacher training	Teachers	Consultants	Unclear	Yes (by coaches)

<b>Study</b>	<b>Intervention name</b>	<b>Intervention category</b>	<b>Who receives the intervention?</b>	<b>Who delivers?</b>	<b>Duration</b>	<b>Fidelity assessed?</b>
Gettinger & Stoiber (2006)	Functional assessment, collaboration, and evidence-based treatment (FACET)	Student intervention	Classroom teacher supported by team of school psychologist, school building principal, and at least one additional special service provider (e.g., social worker, speech and language therapist) all trained to support selected children.	Teacher directly supported by team	2 years	Yes
Hops 1978	CLASS programme	Teacher training	Teachers	consultant trainers	2 months	Unclear
Iovannone 2009	PTR	Student intervention	Teachers	PTR consultants	42 hours	Yes (by coaches/observers)
Kirkhaug 2016	IY TCM	Teacher training	Teachers and after-school staff	Qualified IY TCM group leaders	83 (45 intervention & 38 control)	Yes
Long 2018	GBG & MST	Reward system	Students	Teacher	2 years	Yes (by teacher carrying out intervention)
Palcic 2009	DRC	Student intervention	Students	Teacher	5 weeks	Yes (integrity)
Randolph & Hardage 1973	CBM & group counselling	Teacher training	Teachers and students respectively	Trained professional	31.5 hours	Unclear
Reinke 2014	IY TCM	Teacher training	Teachers	Trained professionals (IYTCM certified leaders)	46 (23 intervention 23 control)	Yes (by coaches)
Simonsen 2011	CICO	Student Intervention	Teachers & students	Teachers	2 years	Yes (by researcher)
Stoiber 2011	PBS	Student intervention	Teachers	Trained professional	5 months	Yes



<b>Study</b>	<b>Intervention name</b>	<b>Intervention category</b>	<b>Who receives the intervention?</b>	<b>Who delivers?</b>	<b>Duration</b>	<b>Fidelity assessed?</b>
Trovato 1992	Teacher-generated/consultant-supported behaviour intervention strategies	Teaching training	Teachers	Trained professional	98 hours	Yes
van den Berg 2018	Classroom seating arrangement	Student placement	Students	N/A	2 years	U
Weinrott 1979	Not specified	Teacher training	Teachers	Not specified	20 teachers/student pairs (10 intervention, 10 control)	Yes (observers)

Table 9. Quality appraisal for Review 2 targeted interventions using RCT design

Study/Criteria	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
Benner 2012	Y	N	Y	N	N	U	Y	U	Y	Y	Y	Y	Y
Breeman 2016	Y	N	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y
Cook 2018a	Y	N	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
Cook 2018b	Y	N	Y	N	N	U	Y	N	Y	Y	Y	Y	Y
Forman 1980	Y	Y	U	N	N	U	Y	N	Y	Y	Y	Y	Y
Forster 2012	U	N	Y	N	N	Y	U	N	Y	Y	Y	Y	Y
Gettinger & Stoiber (2006)	Y	N	Y	U	N	U	N	Y	Y	Y	Y	Y	Y
Hops 1978	Y	N	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y
Iovannone 2009	Y	N	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y
Palcic 2009	U	U	Y	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Reinke 2014	N	N	Y	N	N	N	U	U	Y	Y	Y	Y	Y
Simonsen 2011	Y	N	Y	N	N	N	Y	U	Y	Y	Y	Y	Y
Stoiber 2011	Y	N	Y	N	N	N	Y	U	Y	Y	Y	Y	Y
Trovato 1992	U	N	Y	U	N	N	Y	N	Y	Y	Y	Y	Y
van den Berg 2018	Y	Y	Y	Y	Y	Y	Y	U	Y	Y	Y	Y	Y

Y= Yes, N=No, U=Unclear.

1. Was true randomization used for assignment of participants to treatment groups?
2. Was allocation to treatment groups concealed?
3. Were treatment groups similar at the baseline?
4. Were participants blind to treatment assignment?
5. Were those delivering treatment blind to treatment assignment?
6. Were outcomes assessors blind to treatment assignment?
7. Were treatment groups treated identically other than the intervention of interest?
8. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?
9. Were participants analysed in the groups to which they were randomized?
10. Were outcomes measured in the same way for treatment groups?
11. Were outcomes measured in a reliable way?
12. Was appropriate statistical analysis used?
13. Was the trial design appropriate, and any deviations from the standard RCT design (individual randomization, parallel groups) accounted for in the conduct and analysis of the trial?

Table 10. Quality appraisal for Review 2 targeted interventions with quasi-experimental design

Study	1. Is it clear in the study what is the 'cause' and what is the 'effect' (i.e. there is no confusion about which variable comes first)?	2. Were the participants included in any comparisons similar?	3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	4. Was there a control group?	5. Were there multiple measurements of the outcome both pre and post the intervention/exposure?	6. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?	7. Were the outcomes of participants included in any comparisons measured in the same way?	8. Were outcomes measured in a reliable way?	9. Was appropriate statistical analysis used?
Axberg 2006	Y	Y	Y	Y	Y	Y	Y	Y	Y
Bishop 1996	Y	Y	Y	Y	Y	U	Y	Y	Y
Cook 2017	Y	Y	Y	Y	Y	N	Y	Y	Y
Durlak 1980	Y	Y	Y	Y	Y	Y	Y	Y	Y
Eisenhower 2016	Y	Y	Y	Y	Y	N	Y	Y	Y
Fuchs 1990	Y	Y	Y	Y	Y	Y	Y	Y	Y
Kirkhaug 2016	Y	Y	U	Y	Y	Y	Y	Y	Y
Long 2018	Y	Y	Y	Y	Y	N	Y	Y	Y
Randolph & Hardage 1973	Y	Y	Y	Y	Y	N	Y	Y	Y
Weinrott 1979	Y	U	N	Y	Y	Y	Y	Y	Y

Y= Yes, N=No, U=Unclear.

## Review 2 quantitative synthesis

Review 2 synthesis of the effectiveness of classroom interventions was conducted by categorising interventions. Firstly by universal interventions where intended recipients are all students, and secondly targeted interventions where participants meet certain inclusion criteria to receive the intervention. The latter were typically elevated scores on measures of behavioural difficulties. Universal interventions could be further categorised by teacher training, reward systems, teacher training that involved reward systems or physical activity.

### Universal interventions

Fifteen of the 31 universal intervention studies included data that allowed for the calculation of effect sizes for behaviour outcomes (see table 11). Overall the median effect size was  $d=0.2$ , with an interquartile range quite narrow at 0.33, but indicating that some negative effects were seen. Only a third of these studies reported effects that were both positive and statistically significantly above zero. This ranged from three very large effects and two smaller effects in magnitude, but more precise given large sample sizes. Of those categorised as teacher training, nearly all included a specific reward system as part of this training. A similar median effect size was seen ( $d=0.23$ ), as well as narrow interquartile range (0.263). Seven of these 12 studies were trialling the Incredible Years teacher classroom management intervention: together, these studies showed a smaller effect size for this intervention (median  $d=0.13$ , interquartile range 0.27). The range of effect sizes for this intervention may be due to different samples, outcome measurement, fidelity of intervention delivery and error that would be expected across a range of measures. Of the other interventions that involved teacher training with a reward system, Piwowar (2013) is notable for a large beneficial effect on behaviour ( $d=3.37$ , 95% CI: 2.31 to 4.43), although only 33 teachers participated. This KODEK programme involves a research-based classroom management programme for teachers that provides teachers with a new classroom management strategy which is reviewed as part of the final session. It is notable that this is one of few studies set in a secondary school (in Germany). Although the study quality and confidence interval do not directly call into question this very large effect, the quasi-experimental design, lack of clarity in some elements of reporting and some outcome measures designed by the research team suggest that further trials of this intervention should be encouraged to see if these effects hold.

Finally, three studies of reward systems where effect sizes could be calculated showed differing effects. Two trials of the Good Behaviour Game showed little beneficial effect of the intervention (Humphrey et al, 2018; LeFlot, 2010), whereas Kamps' (2015) trial of Class-Wide Function-Related Intervention Teams (CW-FIT), holds similarities being a group

contingency game that reinforces appropriate behaviour, but demonstrated a very large effect size ( $d=2.00$ , 95% CI: 1.62 to 2.38). Again this effect size is very large and study quality, design (RCT) or confidence interval do not immediately call this into question. However, another included study from the same research team where effect sizes could not be calculated by reviewers indicated more modest, but still beneficial, effects for the same intervention.

Table 11. Effect sizes for Review 2 Universal studies

Author/Date	Intervention name	Intervention Category	Behaviour outcomes	Effect size (d) and 95% confidence interval
Piwowar 2013	KODEK	Teacher training (including reward system)	Munich Observation Inventory & Observation of classroom management	3.37 (2.31 to 4.43)
Fossum 2017	IY TCM	Teacher training, including reward system	Sutter-Eyberg Student Behavior Inventory-Revised; Teacher Report Form; Social Competence and Behavior Evaluation	1.31 (1.16 to 1.46)
Hutchings 2013	IY-TCM	Teacher training, including reward system	Teacher-Pupil Observation Tool	0.51 (-0.64 to 1.66)
Evertson 1989	School-based training programme	Teacher training, including reward system	Observed classroom behaviour	0.33 (-0.40 to 1.06)
Fernandez (2015)	Teacher-Child Interaction Training (TCIT)	Teacher training, including reward system	Teacher rated problem behaviour	0.31 (-0.06 to 0.68)
McGilloway 2010 IYTCM	IYTCM	Teacher training, including reward system	Strengths and Difficulties Questionnaire & Conners Teacher Rating Scale	0.25 (-0.02 to 0.52)
Ford 2018	IY-TCM	Teacher training, including reward system	Strengths and Difficulties Questionnaire	0.13 (0.04 to 0.22)
Murray 2018	IYTCM	Teacher training, including reward system	Revised Teacher Social Competence Scale & Conners Teacher Rating Scale	0.12 (0.01 to 0.23)
Hickey 2017	IY-TCM	Teacher training, including reward system	Strengths and Difficulties Questionnaire & Teacher-Pupil Observation Tool	0.10 (-0.17 to 0.37)
Aasheim 2018	IYTCM	Teacher training, including reward system	Sutter-Eyberg Student Behavior Inventory-Revised & Teacher Report Form	0.03 (-0.10 to 0.16)
Han 2005	RECAP	Teacher training, including reward system	Child Behaviour Checklist & Teacher Report Form	-0.14 (-0.49 to 0.23)
Capella 2012	BRIDGE	Teacher training (no reward system)	Student-Teacher Relationship Scale; School Bullying Experience	0.20 (-0.13 to 0.53)

			Questionnaire; Self-Perception Profile for Children; Behavior Rating Inventory of Executive Function	
Kamps 2015	CW-FIT	Reward system	Observed classroom behaviour	2.00 (1.62 to 2.38)
Leflot 2010	Good Behaviour Game	Reward system	Observed classroom behaviour & peer ratings of problem behaviour	0.09 (-0.14 to 0.32)
Humphrey 2018	GBG	Reward system	Teacher Observation of Classroom Adaptation Checklist	-0.09 (-0.17 to -0.01)

As can be seen in table 11 effect sizes could be calculated for 15 Review 2 universal studies. From table 4 we can see that the remaining studies included in Review 2 are categorised as teacher training including reward systems (5 studies), reward systems (8 studies), teacher training without reward systems (two studies) and physical activity (one study). These categories indicate how the intervention intends to improve behaviour outcomes. The teacher training category represents those interventions that, with the support of coaches or trainers with behavioural expertise, encourage positive student-teacher relations by helping teachers to think of reasons students might misbehave and reflect on how they currently respond. The training can use one or more formats such as on-line, workshops, videos, and one to one or group sessions. The training in these other two studies (Gregory, 2014 and Okonofua, 2016) helped teachers to consider approaches that move away from labelling misbehaving students as troublemakers and exclusionary discipline. The intervention training supports teachers to develop more empathetic and respectful approaches and incorporate them into their teaching in a sustained way. Both these studies suggested by using the skills and knowledge developed in teacher training the rates of exclusionary discipline decreased.

Teacher training including rewards systems category extends the simpler teacher training category described above with the addition of simple, short, feasible and reinforcing rewards or incentives such as tokens that can be exchanged, such as stickers, positive notes, free play, games and physical activity time. These rewards are used as behavioural support with the intention of improving behavioural outcomes, hence as a positive form of contingency management. The five studies (Caldarella et al, 2018; Evertson, 1995; Reinke et al, 2012; Reinke et al, 2018; Thompson et al, 1974) where effect sizes could not be calculated all reported improvements in behavioural outcomes, such as increased on-task engagement, reductions in disruptive and inappropriate behaviour, and increased pro-social behaviour.

The studies in the reward system category have interventions that have components intended to encourage teachers to reinforce behaviours they want to increase such as learning and on-task behaviours. This category of intervention can be distinguished from the teacher training including reward systems category in that the focus here is on the reward system only. All eight studies narratively synthesised demonstrated positive behavioural outcomes, although the positive outcomes in the intervention in Rogeness et al (1977) did not continue into the second year of the study which they suggested was due to staff changes. Also, Spilt et al (2013) found no positive effects for those students with combinations of behavioural and social risks and those from dysfunctional families and suggest these groups may need additional targeted support.

The final review 2 category is classroom physical activity (CPA) which involves students participating in a short period of games or activities as part of their usual academic lesson. This one study (Bartholomew et al, 2018), examined the impact of CPA on the amount of time students were on-task. They found that there was a positive relationship between CPA and time on-task.

#### *Other outcomes*

Table 12 displays the findings for outcomes from Review 2 universal intervention studies that were not characterised as behaviour. The 24 studies reported outcomes other than behaviour with many studies having more than one, e.g. both children and teacher non-behavioural outcomes in a single study. Given the research is focussing on children's behaviour in schools, it is not surprising that the majority of outcomes were in some way focussed on teachers (with 18 distinct outcomes and 27 measures) and children (with 10 distinct outcomes and 17 measures). Other outcomes were quality or fidelity of the intervention (5 and 6 respectively), whole school (both 2), cost (both 2) and with one study a home related non-behavioural outcome and measure. Of the child focussed outcomes 8 studies had attainment as an outcome with the other 9 each a different outcome and measure. For teacher focus, teacher behaviour is the most common outcome in 4 studies and there are a further 16 other distinct outcomes and measures split across either one, two or three studies. While teacher and observer respondents form most of the respondents across the 45 measures, children were clearly asked to respond in 5 measures and parents in two measures.

Due to the quantity of studies and strength of intervention effects, the teacher focused non-behavioural outcomes show the strongest benefits of all, with child focussed interventions having the second most positive effect. As there are small numbers of other outcomes and

measures, we make no overall conclusions on their effects. However, the child focus attainment outcomes had notable positive effects as four studies had strong effects and three had weak effects out of 8 studies in total.



Table 12. Non-behavioural outcomes for Review 2 Universal studies

<b>Author /Date</b>	<b>Intervention name</b>	<b>Intervention Category</b>	<b>Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)</b>	<b>Outcome effect</b>
Aasheim (2018)	IY-TCM	Teacher training (including reward system)	Attainment: academic – T	Small positive effect
Caldarella (2018)	CW-FIT	Teacher training (including reward system)	Social validity - T & C Attainment: academic competence - T	Significant positive effects in both for students at risk of EBD
Capella (2012)	BRIDGE	Teacher training (no reward system)	Relationship: student-teacher – T	A significant positive effect
Dolan (1993)	Good Behaviour Game (GBG)	Reward system	Attainment: reading - C/T (unclear)	No evidence that the GBG improves students' reading
Evertson (1989)	School-based training programme	Teacher training (including reward system)	All teacher behaviour outcomes – O, measures: Instructional management Room arrangement Rules and procedures Meeting student needs Managing student behaviour Classroom climate Organizing activities and physical space Handling student problems during seat work Managing instructional activities Dealing with misbehaviour Monitoring and maintaining accountability Personal Characteristics	Five of 12 ratings were significant. No significant effect Seven of 10 ratings were significant. No clear overall effect. A significant positive effect. A small positive effect Positive effect for preparation component only No significant effect No overall effect No effect Positive effect for keeping students responsible only Small positive effects for teacher as confident and enthusiastic only

... continued

Author /Date	Intervention name	Intervention Category	Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)	Outcome effect
Evertson (1995)	Classroom Organisation and Management Program (COMP)	Teacher training (including reward system)	Attainment: reading comprehension – C & T unclear Attainment: mathematics computation – C & T unclear	Significant gains on reading and maths attainment.
Fernandez (2015)	Teacher-Child Interaction Training (TCIT)	Teacher training (including reward system)	Teacher skill acquisition: teacher verbalisations: - O Teacher satisfaction: with training– T Teacher stress – T	Teachers skills improved and authors suggest teachers were very satisfied with the training and it seemed to help alleviate teacher distress.
Ford (2018)	IY-TCM	Teacher training (including reward system)	Child mental health: difficulties – T & P Other mental health: child difficulties – T & P Other attitude towards school: How I feel about my school – C Other service use: child & adolescent service cost - P	A small improvement in teacher reported children's mental health at 9-months with no effect evident from parent reports. No effect on how children feel about their school. No effect on service use costs
Gregory (2014)	My Teaching Partner-Secondary program (MTP-S)	Teacher training (no reward system)	Teacher sensitivity - O Regard for adolescent perspectives - O Positive climate - O Instructional learning formats: varied use of instructional modalities and strategies - O Analysis and problem solving, engagement in activities that require synthesis, evaluation, and novel application of knowledge – O	Only a broad statement that statistical models showed that the program was beneficial, when accounting for a range of classroom, teacher, and student characteristics
Hickey (2017)	IY-TCM	Teacher training (no reward system)	Other behaviour management: teacher strategies –T Behaviour (teacher & child): teacher-pupil – O	Significant positive effect on teaching strategies. No effect on teacher-pupil interactions
Humphrey (2018)	Good Behaviour Game (GBG)	Reward system	Attainment: reading – O	No evidence it improves reading

Hutchings (2013)	IY-TCM	Teacher training (including reward system)	Behaviour (teacher & child): teacher-pupil - O Other fidelity: teacher satisfaction - T	Significant positive effect No effect given for fidelity
------------------	--------	--------------------------------------------	--------------------------------------------------------------------------------------------	-------------------------------------------------------------

... continued

Author /Date	Intervention name	Intervention Category	Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)	Outcome effect
Kamps (2015)	Class-Wide Function-related Teams Intervention	Reward system	Other teacher behaviour: praise, attention and reprimands - O Other fidelity: procedural fidelity - O	Positive effects – frequency of praise and attention increased, reprimands decreased Intervention implemented with high fidelity
Leflot (2010)	Good Behaviour Game (GBG)	Reward system	Teacher's behaviour management: negative remarks – O	Use of negative remarks reduced.
McGilloway 2010	IY-TCM	Teacher training (including reward system)	Behaviour management: teacher strategies – T Teacher stress – O Home – school collaboration – O Intervention cost: comparison with similar interventions - O	Significant positive effects Significant positive effects Significant positive effects IY-TCM cost is modest in comparison
Murray (2018)	IY-TCM	Teacher training (including reward system)	Academic competence/other learning – T Classroom climate, behaviour management: classroom assessment - O Teacher behaviour: coded impressions - O Other intervention acceptability: teacher satisfaction - T Other SEL, academic: teacher social competence - T Teacher behaviour - T	No significant effect Significant positive effect  Small effect High level of acceptance  Small effect Small effect
Piwowar (2013)	KODEK classroom management competencies"	Teacher training (including reward system)	Behaviour management: classroom management – T Teacher knowledge: classroom management - T Classroom management: competencies – C & O Acceptability – T	Unclear effects. Significant increase in knowledge in 6/8 scores Unclear effects. Significant positive effect

Reinke (2012)	IY-TCM	Teacher training (including reward system)	Attainment - tests of achievement – O	Positive effect
Reinke (2018)	IY-TCM	Teacher training (including reward system)	Attainment: - reading and maths: T Emotional dysregulation - O Social competence – O	No significant effects Significant reduction Significant positive effect
Rogeness (1977)	Behaviour modification & counselling	Reward system	Attainment: science research associate IQ test – T or O unclear Other anxiety: children’s anxiety – T or O unclear	Significant positive effect Significant positive effect

... continued

<b>Author /Date</b>	<b>Intervention name</b>	<b>Intervention Category</b>	<b>Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)</b>	<b>Outcome effect</b>
Spilt (2016)	Good Behaviour Game (GBG)	Reward system	Teacher behaviour: reprimands and praise - O Self-concept: self-perception for children - C Social: school liking and avoidance – C	Small positive effect No effect No effect
Thompson (1974)	Training in behaviour management	Teacher training (including reward system)	Behaviour management: teacher positive and negative events- O	Significant positive effect
Wills (2016)	Class-Wide Function-related Intervention Teams	Reward system	Teacher behaviour: praise and reprimands to individuals and groups – O Classroom management: general and classroom management – O	Significant positive effects (increase in praise and reduction in reprimands)
Wills (2018)	Class-Wide Function-related Intervention Teams	Reward system	Teacher praise – O Teacher reprimand - O Classroom management: general and classroom management – O	Significant positive effect Significant positive effect in 2 of 3 sites Significant positive effects

### Targeted interventions

Twenty out of 25 studies trialling targeted interventions reported statistics that allowed for the calculation of effect sizes for behaviour outcomes. Overall, the median effect size was larger than for the universal interventions ( $d=0.50$ ), but with wide variation across studies (interquartile range 0.96). One might expect the larger median effect size, given the greater need for intervention in the samples and more intensive interventions. More than half of the studies showed statistically significant beneficial effects, although often the confidence intervals were wide on account of the smaller samples for these studies. Of the 11 studies categorised as teacher training, effect sizes were large, but varied from no effect to very large across studies (median  $d=0.58$ , interquartile range 0.87). There was a wider range of different interventions for targeted populations, with only two studies investigating the same intervention (Incredible Years Teacher Classroom Management, although with different populations: children at risk of disruptive behaviour (Reinke et al., 2014) versus children with severe externalising behaviour (Kirkhaug et al., 2016) (small beneficial effect  $d=0.09$  to  $d=0.27$ )). There were mixed effect sizes for those teacher training interventions that involved consultants, ranging from  $d=0.11$  to  $d=1.01$ . Large effect sizes were seen in all three studies trialling different interventions by Cook and colleagues (2017, 2018a, 2018b) that were situated in classrooms with high rates of disruptive and inattentive behaviour. These three interventions were relatively straightforward: Positive greetings at the door, prompts to praise at variable intervals and efforts to improve teacher-student relationships. Study quality and sample size do not appear to call into question these large effects. Perhaps the recruitment of classrooms from schools that had the most disruptive and off-task behaviour might account for the improvement compared to comparison classrooms.

The seven interventions categorised as student interventions showed similar effect sizes as for teacher training (median  $d=0.58$ , interquartile range 0.95). The four interventions recording the larger effects were either functional behavioural assessment interventions involving a wide range of staff or used daily report cards, so all these interventions were to some extent tailored to the behavioural needs of individual children. Although, there is imprecision of some of these large beneficial effects as indicated by confidence intervals; this area therefore warrants further investigation. Finally, as for universal interventions, there was no evidence for benefit of the Good Behaviour Game for either sample ( $d= -.03$ ). These were the only reward systems included that targeted certain populations.

Table 13. Effect sizes for Review 2 Targeted studies

Author/Date	Intervention name	Intervention Category	Behaviour outcomes	Effect size (d) & 95% confidence interval
Cook (2017)	(5):(1)	Teacher training	Behavioral Observation of Students in Schools	1.39 (1.04 to 1.74)
Cook (2018a)	PGD	Teacher training	Behavioral Observation of Students in Schools	1.33 (1.03 to 1.63)
Cook (2018b)	EMR	Teacher training	Behavioral Observation of Students in Schools & Student-Teacher Relationship Scale	1.11 (0.83 to 1.39)
Fuchs (1990)	Behavioural consultation	Teacher training	Revised Behaviour Problem Checklist	1.01 (0.12 to 1.90)
Axberg (2006)	Marte Meo Model	Teacher training	Child Behaviour Checklist & Conners Parent Rating Scale	0.7 (0.01 to 1.39)
Benner (2012)	Behaviour Intervention	Teacher training	Observation classroom behaviour	0.58 (0.09 to 1.07)
Eisenhower (2016)	Starting Strong	Teacher training	Child Behaviour Checklist	0.43 (0.02 to 0.84)
Kirkhaug et al (2016)	IY TCM	Teacher training	Sutter-Eyberg Student Behavior Inventory	0.27 (-0.23 to 0.77)
Hops (1978)	CLASS programme	Teacher training	Observation classroom behaviour	0.11 (-0.75 to 0.97)
Reinke et al (2014)	IY TCM	Teacher training	Teacher-rated problem behaviour	0.09 (-0.49 to 0.67)
Weinrott (1979)	Not specified	Teacher training	Observation classroom behaviour	-0.02 (-0.92 to 0.88)
Palcic (2009)	DRC	Student intervention	Observation classroom behaviour	3.05 (1.98 to 4.12)
Gettinger & Stoiber (2006)	Functional Assessment Collaboration	Student intervention	Observation classroom behaviour	1.37 (0.72 to 2.02)
Stoiber (2011)	PBS	Student intervention	Classroom Competence Observation Form	0.86 (0.31 to 1.41)
Iovannone (2009)	PTR	Student intervention	Social Skills Rating System; Academic Engaged Time	0.58 (0.31 to 0.85)
Simonsen (2011)	Check-in, Check-out	Student intervention	Teacher-rated problem behaviour; Observation classroom behaviour; Social Skills Rating System	0.21 (-0.42 to 0.84)
van den Berg (2018)	Classroom seating arrangement	Student placement	Teacher-rated problem behaviour & peer-rated problem and prosocial behaviour	0.06 (-0.20 to 0.32)
Forster et al (2012)	COMET	Student intervention	Brief Rating of Externalizing Behaviour scale; Conners Teacher Rating Scale; Observation	-0.02 (-0.42 to 0.38)
Breeman (2016)	Good Behaviour Game	Reward system	Teacher-rated problem behaviour	-0.03 (-0.23 to 0.17_)
Long (2018)	Good Behaviour Game	Reward system	Teacher-rated problem behaviour	-0.03 (-0.58 to 0.52)

From table 8 we can see there are six studies that have no calculated effect sizes, four are categorised as teacher training (defined as in the Review 2 universal interventions synthesis above) and two as student intervention. The student intervention category represents interventions that focus training or teaching directly on students whether through counselling or training carried out by school staff, outside experts or a combination of both. Students will have been selected or screened for inappropriate behaviours or emotionally disturbed disorders. The two other studies in this category (Bishop, 1996 and Forman, 1980) demonstrated beneficial behavioural effects such as reducing aggressive and disruptive behaviour or increasing on-task behaviour.

The four other teacher training interventions (Trovato, 1992; Randolph & Hardage, 1973; Hops, 1978; Durlak, 1980) all had beneficial behavioural effects with increased on-task and appropriate behaviours. In addition Hops (1978) found behaviour changes increased significantly over three years of the study with decreasing requirements for students to be referred to special behavioural services and reduced placement of students in special classes. However, Durlak (1980) found behavioural improvements were more significant in students with moderate rather than severe behaviour problems.

#### *Other outcomes*

Table 14 displays the findings for outcomes from Review 2 targeted intervention studies that were not characterised as behaviour. In the 17 studies reporting outcomes other than behaviour there were three areas of focus; teachers (with 7 distinct outcomes and 12 measures), children (9 and 13 respectively) and, quality and fidelity (both 5). When focussing on teachers, relationships were the most common outcome measure with 3 studies. Of the child focussed outcomes 5 studies had attainment as an outcome. While teacher respondents form most of the respondents across the 30 measures, children were asked to respond in 4 measures (all child focussed) and parents in none. The strongest positive effect for any outcome was for the behaviour of teachers outcome, with all three studies reporting significant positive effects; all other outcomes were cumulatively less effective due to their lack of strength and the limited numbers of studies.

Table 14. Non-behavioural outcomes for Review 2 targeted studies.

Author /Date	Intervention name	Intervention Category	Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)	Outcome effect
Axberg (2006)	Marte Meo Model	Teacher training	Behaviour and mental health – T	Clinically significant improvement
Benner (2012)	Behaviour Intervention	Teacher training	Attainment: academic skills – O	Small effect
Bishop (1996)	BTS	Student intervention	Attainment: academic performance - Other (written assessment)	No effect
Breeman (2016)	GBG	Reward system	Relationship: teacher-child relationship – T Social: children's social preference - C Teacher's confidence: sense of self-efficacy - T Feelings/stress: teacher's burnout - T Feelings: child's reflection on teacher-child closeness – C	No effect No effect Effect on engaging students only No effect No effect
Cook (2017)	5:1	Teacher training	Behaviour teachers: ratio of positive-to-negative interactions – O	Significant improvements in positive-to-negative interactions.
Cook (2018)	Establish-maintain-restore (EMR)	Teacher training.	Acceptability: intervention rating – T Fidelity: Implementation -T	Teachers found the PGD strategy to be feasible, reasonable, and acceptable. Fidelity of implementation is unclear
Cook (2018)	Positive greetings at the door (PGD)	Teacher training (including reward system)	Relationship: student-teacher – T Acceptability: intervention rating – T Fidelity: intervention checklist -T	Significant improvement in student-teacher relationships. Adequate levels of acceptability and fidelity
Durlak (1980)	Behavioural treatment and relationship treatment	Teacher training	Behaviour, social: mood – T	No significant effects

... continued



Author /Date	Intervention name	Intervention Category	Outcome category: name of measure: - who the rater was (T=Teacher, C=Child, O=Observer, P=Parent, S=Staff)	Outcome effect
Eisenhower (2016)	Starting Strong	Teacher and parent training	Relationship: student-teacher quality – T Student-teacher: closeness - T Student-teacher: conflict – T	No significant effects in any category.
Forster (2012)	COMET	Student intervention	Behaviour teachers: teacher reprimands - O Behaviour teachers: teacher praise – O	Significant positive effects for both.
Kirkhaug (2016)	IY TCM	Teacher training	Behaviour (closeness and conflict) – T Behaviour, attainment: academic performance – T	No effect Increased academic performance
Long (2018)	GBG & MST	Reward system	Wellbeing/academic: student subjective wellbeing questionnaire – S	No effect
Randolph & Hardage (1973)	DRC Teacher consequences	Student intervention	Attainment: Grade point average - Other (school records) Social: socioeconomic status - O Attainment: school attendance - Other (school records)	Significant positive effect Significant positive effect No effect
Reinke (2014)	CBM & group counselling	Teacher training	Attainment - O	No effect
Simonsen (2011)	IY TCM	Teacher training	Attainment: Academic competence - T	Small effect
Stoiber (2011)	CICO	Reward system	Teacher competency: self-rating - T Teacher competency: accommodating children with challenging behaviour - T Behaviour teachers: ecobehavioural variables – O	Significant positive effects in all outcome measures
van den Berg (2018)	Teacher-generated/consultant-supported behaviour intervention strategies	Teaching	Behaviour/likeability: interpersonal liking - C (children in class) Behaviour and social: group liking - C (whole class)	Positive effects for both, most significant when students sat next to a well-liked and prosocial buddy, or when they were initially disliked

### Review 3 descriptive synthesis

Moving now to review 3, where interventions were focused at a whole school level, twelve RCTs or quasi-experimental studies were included. The summary details of the included studies can be seen in Table 15. Four studies were a mix of whole school and classroom interventions (n=4). The studies were published between 1993 and 2016; they were mainly conducted in the USA (n=7), with none from the UK. The design of the studies included RCTs (n=7), which all included one treatment group, and quasi-experimental studies (n=5) with 1 to 3 treatment groups. Treatment-as-usual was the most frequent comparison type (n=8) and in RCTs randomisation was inevitably always at the level of school. Children were participants in all studies, exclusively so in three, but teachers (n=9), parents (n=1) and all staff (n=5) were also included in some studies. The reported number of recruited participants ranged from 48 to 12344 children, 102 to 1601 teachers, 1211 to 2000 staff, and 4 to 63 schools. Again school level was restricted to the primary level only in most studies (n=9), with middle, secondary or all levels each represented once each. Where reported there were between 7.3% and 53.5% female children in each study sample and only one study reported the mean and standard deviation age at 8.13 years (1.92).

Table 16 provides details of the interventions tested in these studies. Interventions were broadly categorised as either i) school systems for behaviour, ii) multiple interventions taking place within schools, or iii) training for staff. Where treatment time was reported clearly it lasted from 3 to 4 days over 2 years to 7 days. Of the twelve studies, only three failed to report assessment of fidelity.

Table 17 reports study quality for the RCTs (n=5), true randomisation procedures were clearly used for assignment of participants to groups in three studies, but no studies made it clear whether or not allocation to intervention or comparator groups was concealed and in only two studies it was clear that groups were similar at baseline. None of those delivering or assessing the effects of the interventions were blind to group assignment, although this was unclear for observer assessors in one study. Groups were not treated identically other than the intervention of interest in three studies, this was not clear in the remaining two. Follow-up outcomes measures taking place beyond the end of the intervention were not completed in one study, while for three studies differences in number of participants completing post-test measures were unclear. In one study it was not clear that participants were analysed in the groups to which they were randomised. In all studies outcomes were measured in the same way for intervention and comparator groups. In four studies outcomes were measured in a reliable way, for the other this was unclear. All included studies applied appropriate design and analysis; any deviations from the standard RCT design (individual randomization,

parallel groups) accounted for in the conduct and analysis of the trial. All studies applied similar measures to all comparison groups and measures were universally valid and reliable.

Table 18 reports study quality for the quasi-experimental studies (n=7). In four studies, the participants included in any comparisons were similar, while for the remaining three studies this was not clearly reported. The participants included in any comparisons were receiving similar experiences, other than the exposure or intervention of interest in four studies, but this was unclear in two studies. There was a control group and multiple measurements of the outcome both before and after the intervention/exposure in all studies. Follow-up outcome measures were either used and if not, changes in the number of participants completing measures after the intervention period were adequately described and analysed in three studies. This was not the case in the remaining four studies. In all studies the outcomes of participants included in any comparisons were measured in the same way, with the use of valid and reliable measures and with appropriate statistical analysis with the exception of one study where the appropriateness of statistical analysis was not clear.

Table 15. Study characteristics for Review 3 included studies

<b>Study</b>	<b>Country</b>	<b>Design</b>	<b>Type of control</b>	<b>Sample</b>	<b>Sample size</b>	<b>School level</b>	<b>Type of sample</b>
Bodin (2016)	Sweden	Quasi	Treatment as usual	Children	13 schools were assigned to treatment (1,867 youth, 119 teachers) and 10 to the control group (1,340 youth, 69 teachers).	Eligible schools comprised school grades 4–9 (ages 10/11–15/16) in public and private schools level not stated.	All children
Bradshaw 2010	USA	RCT	TAU	School staff & students	37 schools.	Elementary	Whole schools
Bradshaw 2012	USA	RCT	TAU (refrained from implementing SWPBIS for 4 years)	Children only	12344 recruited, 11738 analysed	Elementary	All children
Gottfredson 1993	USA	Quasi	TAU	All staff and students	5719 students	middle schools	All children
Holtzapple 2011	USA	RCT	TAU	Children & teachers	8350 children, 469 teachers	High school	All children
Horner 2009	USA	RCT	Waitlist	School staff & students	63 schools	Elementary	Whole schools
Nelson 1996	USA	Quasi	TAU	Teachers & students	102 teachers baseline. Students: 24 target students &	Elementary	Externalising behaviour and typical students

					24 matched criterion students at baseline.		nominated by teachers
Sorlie 2007	Norway	Quasi	TAU	Children & teachers	780 children & 108 teachers recruited.	Elementary	All children
Sorlie 2015	Norway	Quasi	TAU	Children & teachers	48 schools, 1211 staff members, 5379 children	Primary	All children
Sorlie 2015	Norway	Quasi.	Treatment as usual	entire school staff and children	At baseline (T1), 37 schools with 11,367 students in 1st to 7th grades and 1333 employees.	Primary	All
Ward & Gersten 2013	USA	RCT	Waitlist	Whole school	32 schools. Approximately 2100-2200 staff responses. Approximately 7500 student surveys.	Elementary	All children
Waschbush 2005	Canada	quasi	Treatment as usual	Parents of all children invited to participate	1,115 children their teachers and parents.	Elementary	All children in

Table 16. Intervention details for Review 3 studies

Study	Intervention name	Intervention category	Whole school details	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Bodin (2016)	Prevention in School (PS)	Positive school rules, and teacher and parent training	Training for leaders and all teachers	Children, teachers and parents	Researchers and PS consultants	24 months	No
Bradshaw 2010	SWPBIS	Systems & procedures	<i>universal prevention strategy that aims to alter the school environment</i>	SWPBIS team from each school - 6-10 staff members + administrator. Teams train other staff members.	SWPBIS coach (e.g. school psychologist/guidance counsellor, experience working with SWPBIS)	5 years	Yes - fidelity
Bradshaw 2012, School-Wide Positive Behavioural Interventions & Supports	SWPBIS	Universal prevention. Organisational	School-wide expectations for student behaviour, which are taught to all students and staff.	Intervention schools formed SWPBIS teams, comprising 5 to 6 members (eg, teachers, administrators). School-wide behaviour expectations are taught to all students and staff.	SWPBIS teams (teachers, administrators)	4 school years	Yes - fidelity

Study	Intervention name	Intervention category	Whole school details	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Gottfredson 1993	The program (not named)	Multiple interventions: school-, classroom-, and individual-level	<i>team of teachers and administrators to prepare the school for the program that would begin the next fall.</i>	School year	Not reported	School year	Yes – implementation
Holtzapple 2011 CKH	Capturing Kids Hearts Campus by Design	Teacher & administrator training	Teacher, administrators	1 academic year	Unclear	CKH 3 day teacher training. PC 2 day teacher & administrator training.	Yes - fidelity
Horner 2009	SWPBS	Systems & procedures	School-wide expectations for student behaviour, which are taught to all students and staff.	Teams of staff from each school	Regular state personnel formally trained in SWPBS practices	2 years	Yes - "implementation". Fidelity discussed.

Study	Intervention name	Intervention category	Whole school details	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Nelson 1996	The project	Ecological factors, behavioural guidelines, supervision, disciplinary responses, classroom management, longitudinal programming, focused interventions, reactive strategies	school-wide organizational practices to promote positive social behaviour.	Whole schools	Researchers and PS consultants	~1 school year - "	No
Sorlie 2007	Positive behaviour, interactions and learning environment in school (PALS)	Multi-level, multi-component	The core components were defining and establishing school-wide expectations	Children	Certified PALS consultant	3 years. Post-assessment conducted 20 months after baseline.	No



Study	Intervention name	Intervention category	Whole school details	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Sorlie 2015	School-Wide Positive Behaviour Support (SWPBS/N-PALS)	Multi-level, multi-component	<i>Altering the school environment through evidence-based interventions and inclusive strategies.</i>	Children	Certified N-PALS coach. Schools nominate 3/4 teachers, principal, school psychologist & parent to plan, inform, carry out, monitor & report on the interventions & outcomes.	4 years. Post-assessment conducted after 3 years of implementation.	Yes - fidelity
Sorlie 2015	"Preventing Problem Behavior in School" (PPBS) intervention	a four-day in-service training program for a school's entire staff	see study details	All staff trained	Training sessions were led by the program developers	The PPBS included a 30-hour in-service training program for school staff	Yes

Study	Intervention name	Intervention category	Whole school details	Who receives the intervention?	Who delivers?	Duration	Fidelity assessed?
Ward & Gersten 2013	Safe and Civil Schools (SCS)	Positive behaviour support	A comprehensive, multimedia program that guides staff through the process of designing a positive and proactive school-wide discipline plan.	All school staff (see previous)	SCS consultant	2 years	Yes "implementation"
Waschbush 2005	The Behavior Education Support and Treatment (BEST) School Intervention Program: Schoolwide (SW), Targeted-School (TS), and Targeted-Home (TH) Approaches	Adherence to and tracking of behavioural rules, rewards, feedback and individual behavioural needs, both at school and home	SW states Universal.TH and TS targetted	Teachers supported in all 3 interventions to deliver to students. Additional interventionists (no details) helped teacher and delivered programme in homes.	No details just two staff members/interventionists	Approx training 1 month then interventions for 9 months	Yes

Table 17. Review 3 RCT quality table

Study	Bradshaw (2010)	Bradshaw (2012)	Holtzaple (2011)	Horner (2009)	Homer (2016)	Ward & Gersten (2013)
1. Was true randomization used for assignment of participants to treatment groups?	N Note 1	N	Y	Y	N	Y
2. Was allocation to treatment groups concealed?	U	U	U	U	U	U
3. Were treatment groups similar at the baseline?	Y	U	U	U Note 2	U	Y
4. Were participants blind to treatment assignment?	N	U	N	N	N	N
5. Were those delivering treatment blind to treatment assignment?	N	N	N	N	N	N
6. Were outcomes assessors blind to treatment assignment?	N	N	N Note 3	N	N	N
7. Were treatment groups treated identically other than the intervention of interest?	N Note 4	U	N Note 4	N	Y	U
8. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?	U	Y	U Note 6	N Note 7	Y	U
9. Were participants analysed in the groups to which they were randomized?	U	Y	Y	Y Note 8	Y	Y
10. Were outcomes measured in the same way for treatment groups?	Y	Y	Y	Y	Y	Y
11. Were outcomes measured in a reliable way?	Y	U	Y	Y	N	Y
12. Was appropriate statistical analysis used?	Y	Y	Y	Y	Y	Y
13. Was the trial design appropriate, and any deviations from the standard RCT design (individual randomization, parallel groups) accounted for in the conduct and analysis of the trial?	Y	Y	Y	Y	Y	Y

Y= Yes, N=No, U=Unclear.

Note 1 matched on select baseline demographics.

Note 2 Significant difference in enrolments.

Note 3 Teacher – N; Observer – U.

Note 4 differences in school districts resources and infrastructure.

Note 5 different schools have different discipline policies.

Note 6 Intervention school withdrawn for non-adherence.

Note 7 All schools included in analysis.

Note 8 Some control schools removed due to accessing SWPBS.

Table 18. Review 3 Quasi-experimental quality table

Study	Bodin (2016)	Gottfredson (1993)	Nelson (1996)	Sorlie (2007)	Sorlie & Ogden (2015)	Sorlie et al (2015)	Waschbush (2005)
1. Is it clear in the study what is the 'cause' and what is the 'effect' (i.e. there is no confusion about which variable comes first)?	N	Y	Y	Y	Y	Y	Y
2. Were the participants included in any comparisons similar?	Y	U	Y	U	Y	Y	U
3. Were the participants included in any comparisons receiving similar treatment/care, other than the exposure or intervention of interest?	Y	U	N Note 1	Y	Y	Y	U
4. Was there a control group?	Y	Y	Y	Y	Y	Y	Y
5. Were there multiple measurements of the outcome both pre and post the intervention/exposure?	Y	Y	Y	Y	Y	Y	Y
6. Was follow up complete and if not, were differences between groups in terms of their follow up adequately described and analysed?	Y	N	N	N	N	Y	Y
7. Were the outcomes of participants included in any comparisons measured in the same way?	Y	Y	Y	Y	Y	Y	Y
8. Were outcomes measured in a reliable way?	Y	Y	Y	Y	Y	Y	Y
9. Was appropriate statistical analysis used?	Y	Y	Y	U	Y	Y	Y

Y= Yes, N=No, U=Unclear.

Note 1 Different schools.

### Review 3 quantitative synthesis

Only five of the twelve studies had the data needed to calculate effect sizes for behaviour outcomes. Because these studies include quite different interventions, we have considered these studies separately rather than pooling any effect sizes. The median effect size from these studies is small ( $d=0.12$ ). This fits with the median average effect size for school-based universal prevention programmes reported in a large systematic review (Tanner-Smith et al., 2018). However, Sorlie and Ogden (2007) ( $d=0.29$ , 95% CI: -0.15 to 0.73) and more so Waschbush (2005) ( $d=0.57$ , 95% CI: -0.21 to 1.35) show some promise, although the confidence intervals suggest a lack of precision for the beneficial effects on account of the small samples of teachers completing these behaviour outcome measures. Waschbush (2005) does differ from the other interventions in Review 3 as it incorporates both universal and targeted intervention elements. The universal intervention consisted of delivering behavioural components including school rules and a reward system to all students in the school. The targeted aspect of the intervention consisted of providing additions or modifications to the program to children who did not respond or by a consistent failure to meet the weekly reward goal.

Table 19. Effect sizes for Review 3 studies

Author/Date	Intervention name	Behaviour outcomes	Effect size (d) & 95% confidence interval
Waschbush 2005	The Behavior Education Support and Treatment (BEST) School Intervention Program: Schoolwide (SW)	Teacher-rated problem behaviour	0.57 (-0.21 to 1.35)
Sorlie 2007	PALS	Teacher-rated problem behaviour	0.29 (-0.15 to 0.73)
Bodin 2016	Prevention in School (PS)	Teacher-rated problem behaviour	0.12 (-0.18 to 0.42)
Nelson 1996	The project	Devereux Behaviour Rating Scale	0.01 (-0.60 to 0.62)
Gottfredson 1993	All treatment	Teacher and child-rated problem behaviour	-0.01 (-0.29 to 0.27)

Of the remaining seven studies the two systems and procedures interventions (Bradshaw, 2010 and Horner, 2009) aim to support appropriate behaviours and prevent disruptive behaviours by altering the school environment through improved systems (e.g., discipline, reinforcement, data management) and procedures (e.g., office referral, training, leadership). These improvements are hypothesised to lead to mutually supportive and positive behaviours amongst all the staff and students within the school. Following the intervention both studies reported that discipline referrals were low (although Horner had no experimental control to compare).

A multi-level and multi-component intervention, called School-Wide Positive Behaviour Support (Sorlie, 2015), incorporates the approach used in the systems and procedures category supplemented by teaching of school rules, positive expectations and social skills, and systematic praise and encouragement of positive behaviour (including reward cards), monitoring of student behaviour, school-wide corrections with mild and immediate consequences (response cost), time-limited small-group instruction or training in academic or social topics, individual interventions and support plans, classroom management skills, and parent information and collaboration strategies. Also, those students with similar needs and difficulties receive education in academic or social topics, or positive support. Further, high-risk students receive individualized and functional behaviour support plans that can involve education and family interventions. Significant positive effects resulted in student problem behaviour with reduced numbers of segregated students.

A universal prevention and organisational intervention (Bradshaw, 2012) uses a 3-tiered prevention framework in which 2 levels of selective and indicated programs are implemented. These complement the third tier which is the universal school-wide component used in the systems and procedures interventions. This study provides strong evidence that the intervention reduces levels of disruptive behaviour and concentration problems and increases prosocial behaviours.

A teacher and administrator training with no focus on reward system intervention (Holtzapple, 2011) supports teachers to help students develop and use social skills. It also provides teachers with classroom management strategies and special training to selected staff to act as mentors to teachers. The intervention demonstrated improvements in prosocial behaviours and decreases the number of discipline referrals.

A four-day in-service training program for a school's entire staff intervention (Sorlie, 2015) Preventing Problem Behaviour in School incorporates elements of the Sorlie (2015) School-Wide Positive Behaviour Support intervention but focuses solely at the universal level. All staff are trained in combinations of differentiated evidence-based practices, reinforcement of expected prosocial behaviour, discipline for misbehaviours, good direction and establishing a functional support system. The intervention reduces problem behaviour throughout the school but more effectively outside the classroom. Also, beneficial effects were seen in positive behaviour management practices and to a lesser extent in disciplinary strategies to prevent and manage misbehaviour.

A positive behaviour support intervention (Ward, 2013) consists of a program that helps teachers to develop a school-wide discipline plan that teaches appropriate behaviours and correct misbehaviours. The intervention produced improvements in student behaviour, reduced suspensions which were sustained into the second year of training.

Table 20 displays the findings for outcomes from Review 3 whole school studies that were not characterised as behaviour. For the ten studies reporting outcomes other than behaviour, there were four areas of focus; teachers (with 7 distinct outcomes and 7 measures), children (7 and 9 respectively), whole school (8 and 9), school system quality and fidelity (both 3). Only attainment and school climate were measured in multiple studies. All four attainment measures were positive, suggesting that whole school behaviour interventions may benefit attainment. Teachers are most often the respondents across the measures, although children were asked to respond in nine measures. However, there were no parent reports.

Table 20. Non-behavioural outcomes for Review 3 studies.

Author /Date	Intervention name	Intervention Category	Name of measure/outcome - and who the rater was (T=Teacher, C=Child, O=Observer, P=Parent)	Outcome effect
Bodin (2016)	Prevention in School (PS)	Positive school rules, and teacher and parent training	School rules: knowledge - C & T Reported giving of reward – T Reported receipt of rewards – C Classroom climate: at 12 & 24 months – C & T Being bullied: at 12 & 24 months – C	No significant effects in any outcome measures.
Bradshaw (2010)	SWPBIS	Systems & procedures	Other implementation: school-wide evaluation - O Other behaviour system: effective behaviour support – T Attainment: assessment – C	Across all outcome measures effects greatest among at-risk and high-risk children.
Gottfredson (1993)	All treatment	Multiple interventions: school-, classroom-, and individual-level	Other school climate: respect for students –C Other school climate: clarity of rules - C Other school climate: fairness of rules - C Classroom Environment - T & C	Schools that had high-implementation of the intervention improved most significantly on student report of respect for students, rule clarity, fairness of rules, and environment.
Holtzapple 2011	Capturing kids' hearts	Teacher & administrator training	School climate/other prosocial: personal morality, citizenship – O	Significant positive effects
Horner (2009)	SWPBS	Systems & procedures	School climate: school safety – T;  Attainment: reading – C	Significant effect immediately after training then remained the same. Tentative finding of reading improvements
Nelson (1996)	The project	Ecological factors, behavioural guidelines, supervision, disciplinary responses, classroom management, longitudinal programming, focused interventions, reactive strategies	Other implementation: school-wide evaluation – T Other learning: work habits - T Other learning: social growth - T Attainment: language development - T Attainment: reading - T Attainment: maths - T Behaviour management: number of expulsions, suspensions & emergency removals - O Behaviour management: inventory - T Other teacher stress: effects of stress - T Other intervention acceptability: consumer - T	Significant effect Significant positive effects for work habits, social growth and attainment outcome measures.  Significant positive effect Significant positive effect No effect Teachers were satisfied with the project



... continued

Author /Date	Intervention name	Intervention Category	Name of measure/outcome - and who the rater was (T=Teacher, C=Child, O=Observer, P=Parent)	Outcome effect
Sorlie (2007)	Positive behaviour, interactions and learning environment (PALS)	Multi-level, multi-component	Other learning conditions/environment: - T & C Implementation quality - O Teacher attitudes/beliefs: collective efficacy - T	Small negative effect Significant positive effect Significant positive effect
Sorlie (2015)	SWPBS/N-PALS)	Multi-level, multi-component	Other learning conditions/environment - T Implementation quality: effective behaviour support – T	Greater effects in higher fidelity schools. 75% schools achieved target quality
Ward (2013)	Safe and Civil Schools (SCS)	Positive behaviour support	Other school climate: rules and expectations are clearly defined, a safe and secure environment is provided, formal school safety and student discipline policy exists, raining about school safety/ student discipline policy is provided for staff and discipline policy is enforced consistently – T Attainment: English language arts and maths – T	Significant positive effect for all outcome measures  Significant positive effect (14% in maths and 9% in English language arts)
Waschbush (2005)	Behaviour Education Support and Treatment (BEST)	Adherence to and tracking of behavioural rules, rewards, feedback and individual behavioural needs, both at school and home	Relationships: conflict with teacher – T Relationships: closeness with teacher – T Relationships: dependency on teacher – T Impairment: peer - T Impairment: academic - T Impairment: class behaviour - T	Significant positive effect Significant positive effect Significant positive effect Significant positive effect Significant positive effect (except TS intervention)

## Qualitative Comparative Analysis (QCA) Method

A range of effect sizes were seen across studies included in Review 2 and Review 3, this was the case for studies using interventions which could be categorised in similar ways, e.g. reward programmes and teacher training. Therefore, we undertook some additional work to investigate whether certain components of interventions might better predict beneficial effects on behaviour. In order to further investigate how different combinations of components of interventions aiming to improve school behaviour lead to different outcomes in terms of student behaviour, we undertook qualitative comparative analysis (QCA). We focused on universal interventions, i.e. with samples that represented whole classes or schools, rather than those that focused on targeted groups of students, in the QCA reported here. The research question guiding this analysis was: What components of the universal behaviour interventions reviewed are effective for behaviour outcomes?

Qualitative comparative analysis (QCA) is a method that takes a “case” rather than “variable” approach to analysis. Here a case is an intervention that has been evaluated as part of an included study in Review 2 or Review 3. QCA can identify complex (non-linear and non-additive) causal patterns and is appropriate in situations where there are limited cases and a large number of factors that may explain differences in findings. It is therefore particularly appropriate for systematic reviews of complex interventions where there is heterogeneity that might be explained by a number of intervention or contextual features.

QCA uses set relations and formal Boolean logic to find commonalities between different cases with the same outcome (Rihoux & Ragin, 2008). In the current QCA the outcome is the effectiveness of a case (or intervention) for behaviour outcomes. QCA considers the necessity and sufficiency of conditions for an outcome, with ‘condition’ in this case denoting a particular intervention component or contextual factor. The focus of investigation is not the individual study or intervention trialled, but the different configurations of intervention or contextual conditions that together are responsible for interventions leading (or not) to the effective behaviour outcomes.

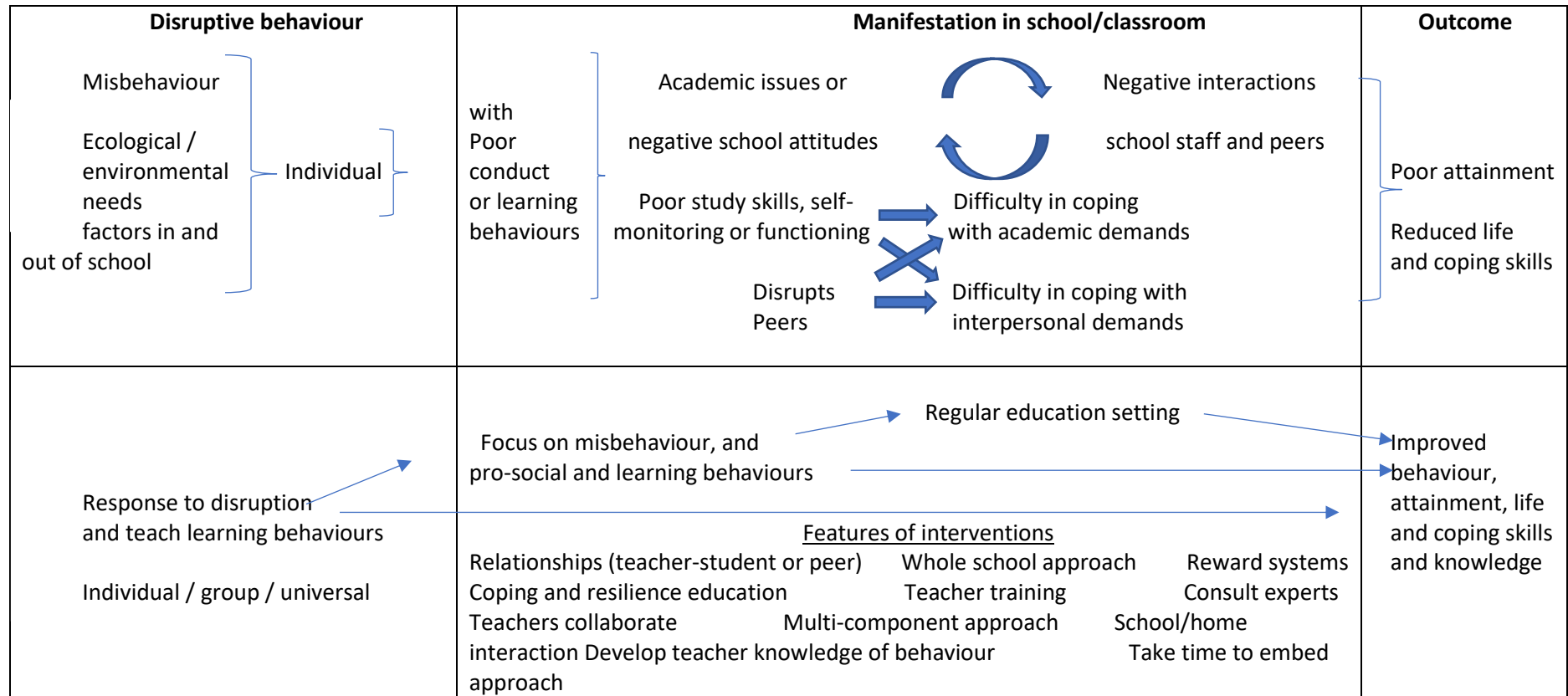
Because QCA is focused on whether the presence or absence of conditions are important to trigger an outcome, a crisp-set QCA analysis sees conditions coded as 1 for present and 0 for absent for each case (Thomas et al., 2014). In fuzzy-set QCA, as will be used here, greater flexibility in categorisation is possible. Here values between 0.5 and 1 are used to

denote membership of a condition or set, and values between 0 and 0.5 used to indicate a spectrum of non-membership (Lee, 2014).

Reviewers began by drawing a logic model that theorised the types of condition that might together impact behaviour outcomes for schoolchildren (see Figure 6). The logic model shows links between underlying causes of student behaviour and manifestation in the school and classroom context that is hypothesised to lead to change in behaviour for pupils. It drew upon Review 1 findings and logic models available for the interventions featuring in studies in Review 2 and 3 (although only Incredible Years Teacher Classroom Management and the Good Behaviour Game had logic models that could be located). We start from the perspective that school staff need to respond to disruptive behaviour and that this is linked to factors in and out of school. Pupil's misbehaviour is seen as a manifestation of unresolved needs. The teacher will become aware of the pupil's poor conduct or learning behaviours, which can be exacerbated when the pupil perceives the reaction from school staff and peers as negative, leading to a vicious circle of academic issues and negative attitudes to school. Poor learning behaviours will manifest as ineffective study skills, self-monitoring or functioning and conduct problems which all can lead to difficulty coping with academic and or emotional demands, resulting in pupil underachievement. In this model longer term outcomes for children on this pathway will be lower achievement and reduced skills and knowledge to meet challenges through their life course.

The interventions we have reviewed are reactive and proactive approaches to prevent or minimise misbehaviour and promote learning behaviour and prosocial behaviour. They range from individualised approaches tailored to individuals to universal responses where all individuals in a class or school receive the same intervention. Some may take place within regular education settings without altering normal teaching and learning. A range of features of interventions might be predicted to improve behaviour. Other outcomes that might be impacted include attainment, coping and life skills.

Figure 6. Logic model to theorise the conditions that may impact outcomes for school pupil behaviour



Identifying specific conditions that might impact on behaviour outcomes according to the logic model and other features of studies included in Review 2 and 3 that previous research predicts would impact behaviour outcomes. Although this initially suggested over 50 conditions, these were reduced according to the theorised importance of the condition in predicting behaviour outcomes, the availability of data in the included studies and the spread of presence and absence of the condition in the included studies (Rihoux 2006). For instance, we were interested in whether a focus on behaviour in regular classroom settings might be a condition seen in effective interventions, but this could not be tested as all but one study analysed here had this focus. The conditions taken forward are indicated in the logic model.

## QCA Findings

For the QCA analysis we followed the steps outlined by Moore and colleagues (2018) in their example of a QCA in a systematic review of interventions for ADHD in school settings.

### *Building the data table*

Twenty-one studies included in Review 2 and Review 3 were used for this QCA, as they met the following criteria: i) included universal rather than targeted samples, ii) included data which allowed effect sizes to be calculated, iii) reported outcomes in terms of student behaviour. Coding of the conditions according to fuzzy set logic was agreed by two reviewers. One reviewer extracted data from original studies and this was checked by the second reviewer with any disagreements resolved through discussion. The codes 0, 0.33, 0.55, 0.67, 1 were used as necessary to refer to partial or full membership of the condition for each case. The effect sizes for academic outcomes were also converted to fuzzified values between 0 and 1 using the calibrate command in the R 3.2.3 software package QCA.

The data extraction and coding gave us a “data table”, i.e. a table consisting of rows that represent the cases (interventions tested in studies) and columns representing the conditions and outcome coded between 0 and 1. The twelve conditions appearing in the data table were whether: 1. Intervention targeted misbehaviour, 2. Intervention targeted prosocial behaviour, 3. Intervention targeted learning behaviour, 4. Intervention targets academic issues, 5. Intervention teaches coping and resilience skills, 6. Intervention is tailored to individual participants, 7. Intervention focuses on improving relationships (with teachers and/or peers), 8. Intervention is applied to whole school, 9. Intervention addresses home/school communication, 10. Intervention includes reward system, 11. Intervention

includes more than 20 hours of teacher training, 12. Intervention includes teachers consulting with experts (see Appendix 4 for data table and criteria for coding).

### *Constructing and checking a 'truth table'*

A truth table takes the data and organises cases by the combinations of causal conditions they exhibit, meaning rows are now specific combinations of the presence or absence of conditions and whether that combination is effective (higher effect sizes for behaviour outcomes). Because four to six conditions would usually be advised (Berg-Schlusser & De Meur, 2009) given the inclusion of 21 cases in the QCA, an iterative process was followed when constructing the truth table and determining the final conditions used. Five conditions appeared in the final truth table: Tailored to individuals, Targets academic issues, coping and resilience skills, relationships, and intensive teacher training (see Table 21).

Table 21. QCA Truth table

Conditions					Outcome			
Tailored to individuals	Targets academic issues	Coping and resilience skills	Relationships	Intensive teaching training	EFFECTIVE	n of cases	sufficiency inclusion score	proportional reduction in inconsistency
0	1	1	0	1	1	1	1	1
1	0	0	1	1	1	1	1	1
0	1	1	0	0	1	1	<b>0.967</b>	<b>0.923</b>
1	1	1	1	1	1	1	<b>0.854</b>	<b>0.706</b>
1	0	0	0	0	0	1	0.724	0.580
0	1	1	1	0	0	1	0.682	0.324
0	0	0	0	1	0	1	0.671	0.507
0	0	0	0	0	0	4	0.470	0.345
0	0	1	1	1	0	10	0.386	0.253

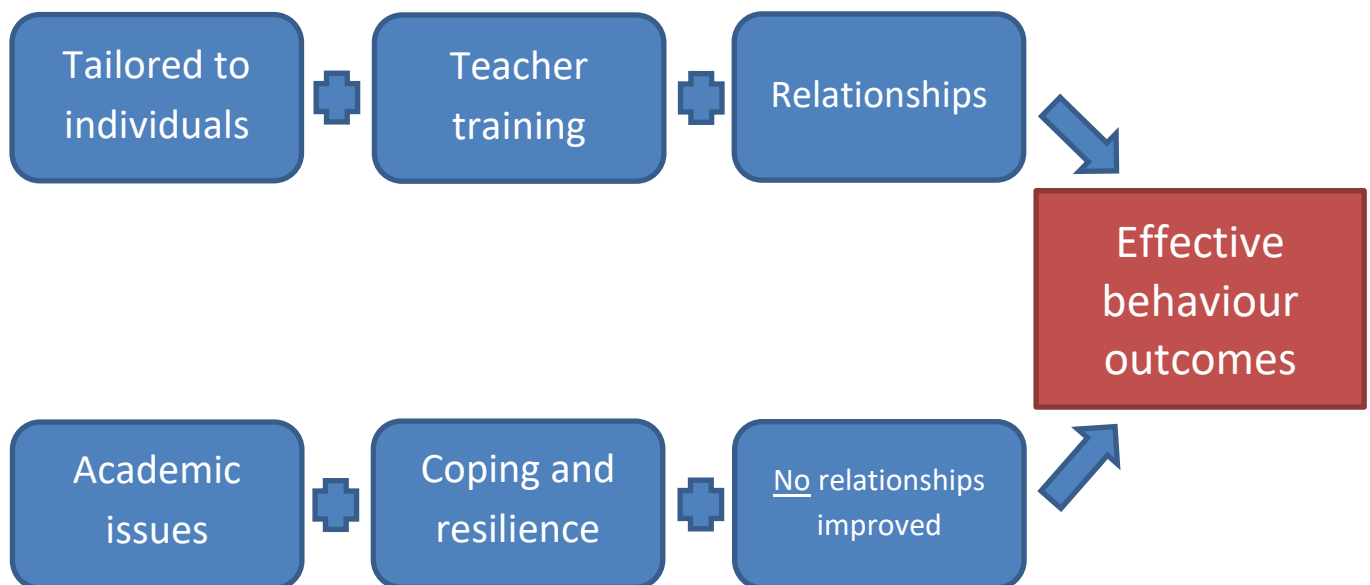
Because targets academic issues and intensive teacher training were often present in configurations that gave effective academic outcomes, we considered whether one or both of these conditions would be enough to clearly explain the causal path to effectiveness, but found that the other three conditions were also important. Indeed, it was not enough to say that the presence of tailored to individuals or more teacher training gave effective outcomes as for some studies the outcome was ineffective when these conditions were present. Note that there are a range of configurations that are effective and ineffective according to the sufficiency inclusion score of 0.8 and that 4 of 21 cases appear in the effective configurations.

Boolean minimisation

This stage aims to simplify the four effective configurations from the data table. The intermediate solution that accounts for remainders – those configurations (n=21) where no cases provide information – and predicts that the presence of the five conditions should lead to effectiveness, as per the logic model, gave two pathways to effectiveness (see Figure 7). The solution coverage of 0.320 indicates the proportion of cases with an effective intervention that fit either pathway. Checks of this model indicated no contradictory configurations and that the model does not also explain ineffective behaviour outcomes, which suggests a good fit for the solution.

Turning to the pathways to effectiveness for behaviour outcomes, the first includes the presence of tailoring to individual intervention recipients, a focus on improving relationships and over 20 hours (i.e. a greater amount than typical) of teacher training. The second pathway indicates that an alternative way of improving behaviour can be to focus on academic issues, teaching coping and resilience skills, but an absence of focus on improving relationships.

Figure 6. Pathways to effectiveness from QCA findings.



#### Logical remainders

Although there were 21 configurations out of a possible 32 that were remainders (meaning no cases provided evidence for effectiveness or ineffectiveness), interventions would be unlikely to fit all these configurations. It would be useful to have further data on configurations where interventions are both tailored to individuals and target academic issues; and both tailored to individuals and teach coping and resilience skills as these only appeared together in one configuration where studies provided data.

## Interpretation

The first causal pathway suggests that tailoring an approach to classroom or school behaviour to individuals and focusing on building relationships can be effective, but this needs to occur with an appropriate amount of teacher training. An alternative is to focus on academic issues and teaching coping and resilience skills, but not focus on relationships. This might be interpreted as a response to behaviour that is more focused on individual student responsibility for behaviour and academic goals. An implication of these findings is a need to identify whether behaviour needs to improve in relation to a child's socio-emotional development (first causal pathway) or a child's academic development (second causal pathway). The behaviour management that might be most effective might be different in each scenario. Therefore a comprehensive behaviour management response needs to consider both aspects of development and evidence suggests this ought to be separate as a focus on improving relationships is only important in response to socio-emotional development. Looking across the four studies that fit the pathways, key commonalities seem to be improving teachers' skills in behaviour management and finding a balance between a whole school approach to behaviour management that also retains the flexibility to respond to individual's needs. Therefore future interventions that combine these approaches would warrant investigation.

## Discussion

### Summary

In Review 1 we reviewed a wide range of research that has provided evidence that certain variables are linked to school behaviour. We located fifty-four articles which were synthesised to produce a model and categories of factors that influence school behaviour. Figure 4 suggests that positive influences are linked to positive behaviour and considering this relationship holds implications for monitoring, not just the reasons why an individual's behaviour might slip, but to anticipate that some influences within and outside school may change behaviour. We categorised the variables that have been shown to influence behaviour in terms of the degree to which schools may impact these factors. Teaching and learning for instance is very much school-based, but aspects like relevance and perceived value of learning can influence behaviour. Home life may be something that can be impacted less directly by school, but parent involvement in learning can be encouraged by schools and awareness of life events or challenges can assist schools in appropriately targeted responses.



In Review 2 we identified 61 studies that assessed the benefits of interventions primarily targeting school behaviour and delivered in classrooms. Thirty-one of these interventions were categorised as universal interventions delivered to all members of classes. Review 2 universal interventions often provided small or large positive effects depending on the study in question. Whether the primary goal of the intervention was to put in place a reward system or not, these interventions trained teachers, most often with at least some benefit for resulting student behaviour. Some interventions trialled in isolated studies show promise and a number of studies trialling the Incredible Years Total Classroom Management programme together show small to medium beneficial effects of this intervention, whereas the effects of the Good Behaviour Game appear negligible. It would be useful to explore the conditions under which Incredible Years is most likely to be effective. It is notable that the included interventions focused on largely positive responses to the challenge of misbehaviour, training teachers to positively encourage learning behaviour and putting in place reward systems, rather than a focus on punitive measures. We did not detect any research investigating interventions that follow zero tolerance-type approaches to managing school behaviour.

In Review 3 we identified 12 studies that assessed the benefits of interventions that primarily targeted behaviour at a whole school level. The whole school approaches often included some features of the Public Health England (2015) framework for a whole school intervention approach, although none contained all eight features depicted in Figure 2. Effect sizes tended to be in line with other school-based universal interventions, although a larger beneficial effect was seen for an intervention that combined both universal as well as targeted elements to a whole school approach to behaviour.

In the qualitative comparative analysis we predicted components of interventions that would improve behaviour outcomes based on our Review 1 findings and previous logic models of included interventions. We then tested these conditions in 21 studies that included universal samples (i.e. child participants were representative of whole school). We found that two different combinations of components of interventions appear to predict effectiveness for behaviour outcomes. Firstly, tailoring to individual intervention recipients, a focus on improving relationships and over 20 hours of teacher training. The second pathway indicates that an alternative way of improving behaviour can be to focus on academic issues, teaching coping and resilience skills, but not to focus on improving relationships.

### Strengths and limitations

Strengths of review 2 and 3 include the use of clear systematic review methods that aid the robustness of the review work and also allow for future updating of the work, not necessarily needing the review team. Review 1 takes an evidence-based approach to attempts to explain school behaviour, when teacher training texts often focus only on addressing behaviour.

Few empirical studies have attempted to identify and theorise the full range of factors that influence school behaviour, therefore Review 1 brought together a 15 year old systematic review that held a similar research question and a range of often primary studies that investigated a small number of variables linked to behaviour rather than taking a more holistic approach. As we searched specifically for frameworks and models, it may be that there are other variables linked to behaviour not recognised in this review, because study authors have not framed the research in an explanatory, framework fashion.

Review 2 and 3 involved a great deal of full text screening. This is indicative of the expansive intervention literature where behaviour might be a secondary outcome. Rather than duplicate previous reviews and not be able to provide detail on the most relevant interventions, we focused very specifically on interventions that primarily targeted improving school behaviour directly, excluding studies that may have hypothesised that behaviour would improve as a result of targeting primary outcomes such as social skills or attainment. However, this means that the review findings must be considered in light of other related work. If, for instance, improving problem solving and thinking skills both improves learning skills and behaviour, this might hold important implications. The findings from Review 3 might suggest that a whole school approach may not suit all students, but equally we had to focus on immediate post-intervention measures, as only one included study (Bodin 2016) recorded longer term follow up measures. Arguably, whole school approaches will need time to bed in to the school community and ethos and therefore follow up outcomes that measure the sustainability of any intervention effects ought to be recommended in future studies.

QCA is limited to using conditions that are reported in included studies. While, we consulted previous literature through Review 1 to theorise how conditions might impact behaviour outcomes, not all conditions that might be relevant will be reported in a journal article write up or have the necessary spread of membership and non-membership of a condition. For instance, we were interested in whether tackling attitudes to school was an effective element of interventions, but this was only present in four cases, so did not have the necessary spread of membership to be tested. Furthermore, QCA holds the limitation that it is essentially analysing whether conditions are present or not and whether interventions are effective or not. Components of interventions or effectiveness is typically more nuanced than

this and by calibrating the range of effect sizes in the 21 studies, effect sizes below 0.3 were considered ineffective (although effect sizes above 0.16 tend to be above the median for school-based universal interventions).

### Suggested recommendations

- A wide range of factors can influence school behaviour. Schools and teachers can only address some of these factors. Staff need to be conscious of some of the factors that may affect behaviour and consider these along with a response to misbehaviour.
- This review has focused on interventions clearly aiming to improve school behaviour, as such a range of studies that assess interventions indirectly improving behaviour are not considered. But a range of high quality reviews can be drawn upon to recommend that behaviour may be improved by focusing on other skills (e.g. problem solving or social and emotional learning) or factors (e.g. parent engagement, symptoms of externalising disorders).
- Either training teachers or putting in place clear reward systems can improve pupil behaviour in the classroom, not just for those pupils most likely to misbehave. A training programme that involves teachers reflecting on their classroom management, trying a new approach and reviewing their progress over time holds promise.
- For schoolchildren who are disruptive, both interventions that train teachers and put in place interventions in the classroom for these individuals can be highly beneficial. It appears that these interventions for targeted populations of students with more behavioural issues are often highly effective when they are tailored to the needs of the individuals involved, rather than attempting to implement the same strategies for all individuals.
- Looking beyond behaviour outcomes, interventions unsurprisingly often led to teachers using more behaviour management techniques. Effects on attainment were measured in some studies and findings seemed to be mixed. However, for whole school approaches to behaviour, there were more consistent beneficial findings in relation to attainment, which might be something to investigate in future research.
- Some relatively straightforward approaches to behaviour management in the classroom have shown very large effect sizes in isolated studies. It would be useful to see if these effects can be replicated in UK settings and of interest to compare such approaches and consider additive effects of different components of behaviour management like teacher-pupil relationships and praise. Likewise many, if not all, teachers recognise the importance of these elements of behaviour management,

however, research appears not to have distinguished what are the key features of effective teacher-pupil relationships and praise.

- We had anticipated stratifying results according to school level, but were surprised to find that amongst 73 studies included in Reviews 2 and 3, only two studies were situated exclusively in high schools or secondary schools. There is a need for further research focused on secondary schools.
- Whole school behaviour programmes can improve behaviour across the student body, but these effects are not always large, which may speak to the time taken to embed a whole school change in behaviour or the difficulty implementing such programmes. One intervention combined a universal and targeted whole school approach, meaning that although a school may have a clear behaviour framework, within this can be flexibility to respond to those students who may struggle.
- It would be worth considering the extent to which whole-school approaches to behaviour interventions fit frameworks for whole school approaches more broadly. Are all staff trained? Is there shared responsibility? Are those in the wider school community involved?
- There are whole school approaches to behaviour management that do not appear to have been studied in randomised controlled trials and quasi-experimental studies. Therefore popular whole school approaches to behaviour management ought to be subject to robust research studies. This may include approaches such as zero tolerance behaviour policies, Teach Like a Champion and Ready to Learn.
- The QCA indicates that effective behaviour management might need to either focus on improving academic and coping skills or train teachers while tailoring approaches to individual students and focusing on improving relationships. This suggests the need to both focus on the more typical management of behaviour through equipping teachers with necessary skills and strategies, tailoring the approach to individual needs and improving relationships with teachers and peers. These approaches can operate both within classes and across the school, as well as acknowledging that when students have the skills to cope and achieve in the classroom behaviour is likely to improve.
- Finally, data linkage from trials to national databases would be helpful to be able to explore in more detail what factors may affect intervention effectiveness and indeed to test some of the ideas stemming from Review 1. More importantly, it would extend the duration of follow up and permit the study of real world outcomes, such as attainment, attendance and exclusions. This would have enabled some of the initial aims of the project to be realised such as which approaches to behaviour are most

effective for pupils with special educational needs and to do a more thorough analysis of which approaches are most effective for pupils with more challenging behaviours.

## References

### Included in Review 1

- Adolphus, K., Lawton, C. L., & Dye, L. (2013). The effects of breakfast on behavior and academic performance in children and adolescents. *Frontiers in human neuroscience*, 7, 425.
- Bear, G. G., & Rys, G. S. (1994). Moral reasoning, classroom behavior, and sociometric status among elementary school children. *Developmental Psychology*, 30(5), 633.
- Bidell, M. P., & Deacon, R. E. (2010). School Counselors Connecting the Dots between Disruptive Classroom Behavior and Youth Self-Concept. *Journal of School Counseling*, 8(9), n9.
- Bitsika, V. (2003). "But I'm Not Really Bad": Using an Ideographic Versus a Nomothetic Approach to Understand the Reasons for Difficult Behaviour in Children. *Journal of Psychologists and Counsellors in Schools*, 13(1), 87-98.
- Boon, H. J. (2011). School moves, coping, and achievement: models of possible interactions. *The Journal of Educational Research*, 104(1), 54-70.
- Borders, A., Earleywine, M., & Huey, S. J. (2004). Predicting problem behaviors with multiple expectancies: Expanding expectancy-value theory. *Adolescence*, 39(155), 539-551.
- Cadioux, A. (2003). A 3-year longitudinal study of self-concept and classroom behavior of grade 1 retained pupils. *Perceptual and motor skills*, 96(2), 371-378.
- Caughy, M. O. B., Nettles, S. M., & O'Campo, P. J. (2007). Community influences on adjustment in first grade: An examination of an integrated process model. *Journal of Child and Family Studies*, 16(6), 819-836.
- Cheung, Y. W. (1997). Family, school, peer, and media predictors of adolescent deviant behavior in Hong Kong. *Journal of Youth and Adolescence*, 26(5), 569-596.
- Chirinos, D. S. (2017). *Examining Dimensions of Expectancy-Value Theory as Predictors of US Latino High School Students' Academic Behaviors and Mathematics Performance* (Doctoral dissertation, George Mason University).
- Cornwall, J. O. H. N. (2015). The Human Element in Education: Nurture, Self-Efficacy and the Psychology of Academic Inclusion. *Innovative Practice and Interventions for Children and Adolescents with Psychosocial Difficulties and Disorders*, 60-82.
- Daniels, A., & Williams, H. (2000). Reducing the need for exclusions and statements for behaviour: the Framework for Intervention Part I. *Educational Psychology in Practice*, 15(4), 220-227.
- Dever, B. V. (2016). Using the expectancy-value theory of motivation to predict behavioral and emotional risk among high school students. *School Psychology Review*, 45(4), 417-433.
- Di Maggio, R., Zappulla, C., & Pace, U. (2016). The relationship between emotion knowledge, emotion regulation and adjustment in preschoolers: A mediation model. *Journal of Child and Family Studies*, 25(8), 2626-2635.
- Dinh, K. T., Roosa, M. W., Tein, J. Y., & Lopez, V. A. (2002). The relationship between acculturation and problem behavior proneness in a Hispanic youth sample: A longitudinal mediation model. *Journal of Abnormal Child Psychology*, 30(3), 295-309.
- DiStefano, C., & Kamphaus, R. W. (2008). Latent growth curve modeling of child behavior in elementary school. *Research in the Schools*, 15(1), 27.
- Ellis, S., & Tod, J. (2018). *Behaviour for Learning: Promoting Positive Relationships in the Classroom*. London: Routledge.
- Eve, R. A. (1978). A study of the efficacy and interactions of several theories for explaining rebelliousness among high school students. *J. Crim. L. & Criminology*, 69, 115.
- Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The "why's" of class size: Student behavior in small classes. *Review of Educational Research*, 73(3), 321-368.
- Garner, P. (2011). *Promoting the conditions for positive behaviour, to help every child succeed*. National College for School Leadership.
- Gottfredson, D. C., Fink, C. M., & Graham, N. (1994). Grade retention and problem behavior. *American Educational Research Journal*, 31(4), 761-784.
- Güleç, S., & Balçık, E. G. (2011). UNDESIRABLE BEHAVIORS ELEMENTARY SCHOOL CLASSROOM TEACHERS ENCOUNTER IN THE CLASSROOM AND THEIR REASONS. *Bulgarian Journal of Science & Education Policy*, 5(2).
- Hastings, R. P. (2005). Staff in special education settings and behaviour problems: Towards a framework for research and practice. *Educational Psychology*, 25(2-3), 207-221.
- Haynes, N. M. (1990). Influence of self-concept on school adjustment among middle-school students. *The Journal of Social Psychology*, 130(2), 199-207.

- Heaven, P. C., Leeson, P., & Ciarrochi, J. (2009). Personality development at school: Assessing a reciprocal influence model of teachers' evaluations and students' personality. *Journal of Research in Personality, 43*(5), 815-821.
- Hershfeldt, P. A., Sechrest, R., Pell, K. L., Rosenberg, M. S., Bradshaw, C. P., & Leaf, P. J. (2009). Double-Check: A framework of cultural responsiveness applied to classroom behavior. *Teaching Exceptional Children Plus, 6*(2), 2-18.
- Jackson, Y., & Frick, P. J. (1998). Negative life events and the adjustment of school-age children: Testing protective models. *Journal of Clinical Child Psychology, 27*(4), 370-380.
- Kemp, D. E., & Center, D. B. (2003). An investigation of Eysenck's Antisocial Behavior Hypothesis in general education students and students with behavior disorders. *Personality and Individual Differences, 35*(6), 1359-1371.
- Kerr, M. A., Black, M. M., & Krishnakumar, A. (2000). Failure-to-thrive, maltreatment and the behavior and development of 6-year-old children from low-income, urban families: a cumulative risk model. *Child abuse & neglect, 24*(5), 587-598.
- Knutson, J. F., DeGarmo, D. S., & Reid, J. B. (2004). Social disadvantage and neglectful parenting as precursors to the development of antisocial and aggressive child behavior: Testing a theoretical model. *Aggressive behavior, 30*(3), 187-205.
- LaRoque, P. T. (2007). *Peer Group Processing Model: Adolescent peer group orientation and its influence on school conduct*. University of Wisconsin--Madison.
- Liao, Y., Shonkoff, E. T., Barnett, E., Wen, C. F., Miller, K. A., & Eddy, J. M. (2015). Brief report: Examining children's disruptive behavior in the wake of trauma—A two-piece growth curve model before and after a school shooting. *Journal of adolescence, 44*, 219-223.
- Liasidou, A. (2016). Discourse, power interplays and 'disordered identities': an intersectional framework for analysis and policy development. *Emotional and Behavioural Difficulties, 21*(2), 228-240.
- Lochman, J. E., & Wells, K. C. (2002). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the Coping Power program. *Development and psychopathology, 14*(4), 945-967.
- Lopez, C., & DuBois, D. L. (2005). Peer victimization and rejection: Investigation of an integrative model of effects on emotional, behavioral, and academic adjustment in early adolescence. *Journal of Clinical Child and Adolescent Psychology, 34*(1), 25-36.
- Lorber, M. F., & Egeland, B. (2011). Parenting and infant difficulty: Testing a mutual exacerbation hypothesis to predict early onset conduct problems. *Child Development, 82*(6), 2006-2020.
- Lyons, C. W., & O'Connor, F. (2006). Constructing an integrated model of the nature of challenging behaviour: a starting point for intervention. *Emotional and behavioural difficulties, 11*(3), 217-232.
- McDermott, P. A., Mordell, M., & Stoltzfus, J. C. (2001). The organization of student performance in American schools: Discipline, motivation, verbal learning, nonverbal learning. *Journal of Educational Psychology, 93*(1), 65.
- Mooij, T. (1999). Promoting prosocial pupil behaviour: 1-A multilevel theoretical model. *British Journal of Educational Psychology, 69*(4), 469-478.
- Nelson, B., Martin, R. P., Hodge, S., Havill, V., & Kamphaus, R. (1999). Modeling the prediction of elementary school adjustment from preschool temperament. *Personality and individual differences, 26*(4), 687-700.
- Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology, 34*(3), 185-194.
- Nolan, K. M. (2011). Oppositional behavior in urban schooling: Toward a theory of resistance for new times. *International Journal of Qualitative Studies in Education, 24*(5), 559-572.
- Norwich, B., & Rovoli, I. (1993). Affective factors and learning behaviour in secondary school mathematics and English lessons for average and low attainers. *British Journal of Educational Psychology, 63*(2), 308-321.
- Núñez, J. L., & León, J. (2015). Autonomy support in the classroom. *European Psychologist*.
- Olvera, P. (2007). *Assessment of Antisocial Behavior in Elementary School Children: Manifestations, Risk Factors, Models of Development and Assessment of Comorbid Learning Disabilities*. ProQuest.
- Pingault, J. B., Tremblay, R. E., Vitaro, F., Japel, C., Boivin, M., & Côté, S. M. (2015). Early nonparental care and social behavior in elementary school: Support for a social group adaptation hypothesis. *Child development, 86*(5), 1469-1488.
- Powell, S., & Tod, J. (2004). *A systematic review of how theories explain learning behaviour in school contexts*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

- Richards, H. C. (1989). *Moral Reasoning and Classroom Conduct: A Closer Look*. Paper presented at the Biennial Meeting of the Society for Research in Child Development (Kansas City, MO, April 27-30, 1989).
- Richards, H. C. (1984). *Moral Reasoning and Classroom Conduct: A Replication*. Paper presented at the Annual Convention of the Psychological Association (92nd, Toronto, Ontario, Canada, August 24-28, 1984).
- Slee, R. (2014). Evolving theories of student disengagement: a new job for Durkheim's children?. *Oxford Review of Education*, 40(4), 446-465.
- Smith, G. (2010). *Relation of Parental Social, Cognitive, and Behavioral Frameworks With School Behavior of High School Students*. Doctor of Education in Educational Leadership Thesis. University of Phoenix.
- Sullivan, C. J., & Hirschfield, P. (2017). Problem behavior in the middle school years: An assessment of the social development model. In *Developmental and Life-course Criminological Theories* (pp. 173-200). Routledge.
- Svedhem, L. (1994). Social network and behaviour problems among 11-13-year-old schoolchildren. A theoretical and empirical basis for network therapy. *Acta psychiatrica Scandinavica. Supplementum*, 381, 1-84.
- Tremblay, R. E. (2010). Developmental origins of disruptive behaviour problems: the 'original sin' hypothesis, epigenetics and their consequences for prevention. *Journal of Child Psychology and Psychiatry*, 51(4), 341-367.

#### Included in Review 2

##### Universal Interventions

- Aasheim, M., Reedtz, C., Handegård, B. H., Martinussen, M., & Mørch, W. T. (2018). Evaluation of the Incredible Years Teacher Classroom Management Program in a Regular Norwegian School Setting. *Scandinavian Journal of Educational Research*, 1-14.
- Bartholomew, J. B., Golaszewski, N. M., Jowers, E., Korinek, E., Roberts, G., Fall, A., & Vaughn, S. (2018). Active learning improves on-task behaviors in 4th grade children. *Preventive medicine*, 111, 49-54.
- Caldarella, P., Larsen, R. A., Williams, L., Wills, H., Kamps, D., & Wehby, J. H. (2018). Effects of CW-FIT on Teachers' Ratings of Elementary School Students at Risk for Emotional and Behavioral Disorders. *Journal of Positive Behavior Interventions*, 20(2), 78-89.
- Cappella, E., Hamre, B. K., Kim, H. Y., Henry, D. B., Frazier, S. L., Atkins, M. S., & Schoenwald, S. K. (2012). Teacher consultation and coaching within mental health practice: Classroom and child effects in urban elementary schools. *Journal of consulting and clinical psychology*, 80(4), 597.
- Dolan, L. J., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., ... & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. *Journal of Applied Developmental Psychology*, 14(3), 317-345.
- Evertson, C. M. (1989). Improving elementary classroom management: A school-based training program for beginning the year. *The Journal of Educational Research*, 83(2), 82-90.
- Evertson, C. M. (1995). Classroom Organization and Management Program. Revalidation Submission to the Program Effectiveness Panel, US Department of Education.
- Fernandez, M. A., Adelstein, J. S., Miller, S. P., Areizaga, M. J., Gold, D. C., Sanchez, A. L., ... & Gudiño, O. G. (2015). Teacher-Child Interaction Training: A pilot study with random assignment. *Behavior therapy*, 46(4), 463-477.
- Ford, T., Hayes, R., Byford, S., Edwards, V., Fletcher, M., Logan, S., ... & Ganguli, P. (2018). The effectiveness and cost-effectiveness of the Incredible Years® Teacher Classroom Management programme in primary school children: results of the STARS cluster randomised controlled trial. *Psychological medicine*, 1-15.
- Fossum, S., Handegård, B. H., & Drugli, M. B. (2017). The Incredible Years teacher classroom management programme in kindergartens: Effects of a universal preventive effort. *Journal of Child and Family Studies*, 26(8), 2215-2223.
- Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. C. (2014). Effects of a professional development program on behavioral engagement of students in middle and high school. *Psychology in the Schools*, 51(2), 143-163.
- Han, S. S., Catron, T., Weiss, B., & Marciel, K. K. (2005). A teacher-consultation approach to social skills training for pre-kindergarten children: Treatment model and short-term outcome effects. *Journal of Abnormal Child Psychology*, 33(6), 681-693.



- Hickey, G., McGilloway, S., Hyland, L., Leckey, Y., Kelly, P., Bywater, T., ... & O'Neill, D. (2017). Exploring the effects of a universal classroom management training programme on teacher and child behaviour: A group randomised controlled trial and cost analysis. *Journal of Early Childhood Research*, 15(2), 174-194.
- Homer, R., Hew, K. F., & Tan, C. Y. (2018). Comparing Digital Badges-and-Points with Classroom Token Systems: Effects on Elementary School ESL Students' Classroom Behavior and English Learning. *Journal of Educational Technology & Society*, 21(1), 137-151.
- Humphrey, N., Hennessey, A., Ashworth, E., Frearson, K., Black, L., Petersen, K., Wo, L., Panayiotou, M., Lendrum, A., Wigelsworth, M., Birchinnall, L., Squires, G. & Pampaka, M.
- Hutchings, J., Martin-Forbes, P., Daley, D., & Williams, M. E. (2013). A randomized controlled trial of the impact of a teacher classroom management program on the classroom behavior of children with and without behavior problems. *Journal of School Psychology*, 51(5), 571-585.
- Kamps, D., Wills, H., Dawson-Bannister, H., Heitzman-Powell, L., Kottwitz, E., Hansen, B., & Fleming, K. (2015). Class-wide function-related intervention teams "CW-FIT" efficacy trial outcomes. *Journal of positive behavior interventions*, 17(3), 134-145.
- Leflot, G., van Lier, P. A., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of abnormal child psychology*, 38(6), 869-882.
- McGilloway, S., Hyland, L., Mhaille, G., Lodge, A., O'Neill, D., Kelly, P., ... & Donnelly, M. (2011). Positive classrooms, positive children: A randomised controlled trial to investigate the effectiveness of the incredible years teacher classroom management programme in an Irish context (short-term outcomes). *Ireland: Archways*.
- Murray, D. W., Rabiner, D. L., Kuhn, L., Pan, Y., & Sabet, R. F. (2018). Investigating teacher and student effects of the Incredible Years Classroom Management Program in early elementary school. *Journal of school psychology*, 67, 119-133.
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). Brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences*, 113(19), 5221-5226.
- Piwowar, V., Thiel, F., & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, 30, 1-12.
- Reinke, W. M., Herman, K. C., & Dong, N. (2018). The Incredible Years teacher classroom management program: Outcomes from a group randomized trial. *Prevention Science*, 19(8), 1043-1054.
- Reinke, W. M., Stormont, M., Webster-Stratton, C., Newcomer, L. L., & Herman, K. C. (2012). The Incredible Years Teacher Classroom Management program: using coaching to support generalization to real-world classroom settings. *Psychology in the Schools*, 49(5), 416-428.
- Rogeness, G. A., Stokes, J. P., Bednar, R. A., & Gorman, B. L. (1977). School intervention program to increase behaviors and attitudes that promote learning. *Journal of community psychology*, 5(3), 246-256.
- Spilt, J. L., Koot, J. M., & van Lier, P. A. (2013). For whom does it work? Subgroup differences in the effects of a school-based universal prevention program. *Prevention Science*, 14(5), 479-488.
- Spilt, J. L., Leflot, G., Onghena, P., & Colpin, H. (2016). Use of praise and reprimands as critical ingredients of teacher behavior management: Effects on children's development in the context of a teacher-mediated classroom intervention. *Prevention Science*, 17(6), 732-742.
- Thompson, M., Brassell, W. R., Persons, S., Tucker, R., & Rollins, H. (1974). Contingency management in the schools: How often and how well does it work?. *American Educational Research Journal*, 11(1), 19-28.
- Wills, H., Kamps, D., Caldarella, P., Wehby, J., & Romine, R. S. (2018). Class-wide Function-Related Intervention Teams (CW-FIT): Student and Teacher Outcomes from a Multisite Randomized Replication Trial. *The Elementary School Journal*, 119(1), 29-51.
- Wills, H., Kamps, D., Fleming, K., & Hansen, B. (2016). Student and teacher outcomes of the class-wide function-related intervention team efficacy trial. *Exceptional Children*, 83(1), 58-76.
- Wills, H., Wehby, J., Caldarella, P., Kamps, D., & Swinburne Romine, R. (2018). Classroom Management That Works: A Replication Trial of the CW-FIT Program. *Exceptional Children*, 0014402918771321.
- EEF (2018). *Good Behaviour Game Evaluation Report and Executive Summary*. London: Education Endowment Foundation.

### Targeted Interventions

- Axberg, U., Hansson, K., Broberg, A. G., & Wirtberg, I. (2006). The Development of a Systemic School-Based Intervention: Marte Meo and Coordination Meetings. *Family process*, *45*(3), 375-389.
- Benner, G. J., Nelson, J. R., Sanders, E. A., & Ralston, N. C. (2012). Behavior intervention for students with externalizing behavior problems: Primary-level standard protocol. *Exceptional Children*, *78*(2), 181-198.
- Bishop, G. B., Rosen, L. A., Miller, C. D., & Hendrickson, J. (1996). Evaluation of the Boy's Town motivation system in a US school setting. *School Psychology International*, *17*(2), 125-131.
- Breeman, L. D., van Lier, P. A., Wubbels, T., Verhulst, F. C., van der Ende, J., Maras, A., ... & Tick, N. T. (2016). Effects of the Good Behavior Game on the behavioral, emotional, and social problems of children with psychiatric disorders in special education settings. *Journal of Positive Behavior Interventions*, *18*(3), 156-167.
- Cook, C. R., Coco, S., Zhang, Y., Fiat, A. E., Duong, M. T., Renshaw, T. L., ... & Frank, S. (2018). Cultivating Positive Teacher–Student Relationships: Preliminary Evaluation of the Establish–Maintain–Restore (EMR) Method. *School Psychology Review*, *47*(3), 226-243.
- Cook, C. R., Fiat, A., Larson, M., Daikos, C., Slemrod, T., Holland, E. A., ... & Renshaw, T. (2018). Positive Greetings at the Door: Evaluation of a Low-Cost, High-Yield Proactive Classroom Management Strategy. *Journal of Positive Behavior Interventions*, 1098300717753831.
- Cook, C. R., Grady, E. A., Long, A. C., Renshaw, T., Coddling, R. S., Fiat, A., & Larson, M. (2017). Evaluating the impact of increasing general education teachers' ratio of positive-to-negative interactions on students' classroom behavior. *Journal of Positive Behavior Interventions*, *19*(2), 67-77.
- Durlak, J. A. (1980). Comparative effectiveness of behavioral and relationship group treatment in the secondary prevention of school maladjustment. *American Journal of Community Psychology*, *8*(3), 327-339.
- Eisenhower, A., Taylor, H., & Baker, B. L. (2016). Starting strong: A school-based indicated prevention program during the transition to kindergarten. *School Psychology Review*, *45*(2), 141-170.
- Forman, S. G. (1980). A comparison of cognitive training and response cost procedures in modifying aggressive behavior of elementary school children. *Behavior therapy*.
- Forster, M., Sundell, K., Morris, R. J., Karlberg, M., & Melin, L. (2012). A randomized controlled trial of a standardized behavior management intervention for students with externalizing behavior. *Journal of Emotional and Behavioral Disorders*, *20*(3), 169-183.
- Fuchs, D., Fuchs, L. S., Bahr, M. W., Fernstrom, P., & Stecker, P. M. (1990). Prereferral intervention: A prescriptive approach. *Exceptional Children*, *56*(6), 493-513.
- Gettinger, M., & Stoiber, K. C. (2006). Functional assessment, collaboration, and evidence-based treatment: Analysis of a team approach for addressing challenging behaviors in young children. *Journal of School Psychology*, *44*(3), 231-252.
- Hops, H., Walker, H. M., Fleischman, D. H., Nagoshi, J. T., Omura, R. T., Skindrud, K., & Taylor, J. (1978). CLASS: A standardized in-class program for acting-out children: II. Field test evaluations. *Journal of Educational Psychology*, *70*(4), 636.
- Iovannone, R., Greenbaum, P. E., Wang, W., Kincaid, D., Dunlap, G., & Strain, P. (2009). Randomized Controlled Trial of the Prevent—Teach—Reinforce (PTR) Tertiary Intervention for Students With Problem Behaviors: Preliminary Outcomes. *Journal of emotional and behavioral disorders*, *17*(4), 213-225.
- Kirkhaug, B., Drugli, M. B., Handegård, B. H., Lydersen, S., Åsheim, M., & Fossum, S. (2016). Does the Incredible Years Teacher Classroom Management Training programme have positive effects for young children exhibiting severe externalizing problems in school?: a quasi-experimental pre-post study. *BMC psychiatry*, *16*(1), 362.
- Long, A. C., Renshaw, T. L., & Camarota, D. (2018). Classroom Management in an Urban, Alternative School: a Comparison of Mindfulness and Behavioral Approaches. *Contemporary School Psychology*, *22*(3), 233-248.
- Palcic, J. L., Jurbergs, N., & Kelley, M. L. (2009). A comparison of teacher and parent delivered consequences: Improving classroom behavior in low-income children with ADHD. *Child & Family Behavior Therapy*, *31*(2), 117-133.
- Randolph, D. L., & Hardage, N. C. (1973). Behavioral consultation and group counseling with potential dropouts. *Elementary School Guidance & Counseling*.
- Reinke, W. M., Stormont, M., Herman, K. C., Wang, Z., Newcomer, L., & King, K. (2014). Use of coaching and behavior support planning for students with disruptive behavior within a universal classroom management program. *Journal of Emotional and Behavioral Disorders*, *22*(2), 74-82.

- Simonsen, B., Myers, D., & Briere III, D. E. (2011). Comparing a behavioral check-in/check-out (CICO) intervention to standard practice in an urban middle school setting using an experimental group design. *Journal of Positive Behavior Interventions*, 13(1), 31-48.
- Stoiber, K. C., & Gettinger, M. (2011). Functional assessment and positive support strategies for promoting resilience: Effects on teachers and high-risk children. *Psychology in the Schools*, 48(7), 686-706.
- Trovato, J., Harris, J., Pryor, C. W., & Wilkinson, S. C. (1992). Teachers in regular classrooms: An applied setting for successful behavior programming. *Psychology in the Schools*, 29(1), 52-61.
- Van den Berg, Y. H. M., & Stoltz, S. E. M. J. (2018). Enhancing social inclusion of children with externalizing problems through classroom seating arrangements: a randomized controlled trial. *Journal of emotional and behavioral disorders*, 26(1), 31-41.
- Walker, H. M., Kavanagh, K., Stiller, B., Golly, A., Severson, H. H., & Feil, E. G. (1998). First step to success: An early intervention approach for preventing school antisocial behavior. *Journal of emotional and behavioral disorders*, 6(2), 66-80.
- Weinrott, M. R. (1979). The utility of tracking child behavior. *Journal of abnormal child psychology*, 7(3), 275-286.

#### Included in Review 3

- Bodin, M. C., South, S. H., & Ingemarson, M. (2016). A Quasi-Randomized Trial of a School-Wide Universal Prevention Program: Results and Lessons Learned. *Scandinavian Journal of Educational Research*, 60(4), 449-476.
- Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, 12(3), 133-148.
- Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, peds-2012.
- Gottfredson, D. C., Gottfredson, G. D., & Hybl, L. G. (1993). Managing adolescent behavior a multiyear, multischool study. *American Educational Research Journal*, 30(1), 179-215.
- Holtzapple, C. K., Griswold, J. S., Cirillo, K., Rosebrock, J., Nouza, N., & Berry, C. (2011). Implementation of a school-wide adolescent character education and prevention program: Evaluating the relationships between principal support, faculty implementation, and student outcomes. *Journal of Research in Character Education*, 9(1), 71.
- Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, 11(3), 133-144.
- Nelson, J. R. (1996). Designing schools to meet the needs of students who exhibit disruptive behavior. *Journal of Emotional and Behavioral Disorders*, 4(3), 147-161.
- Sørli, M. A., & Ogden, T. (2007). Immediate Impacts of PALS: A school-wide multi-level programme targeting behaviour problems in elementary school. *Scandinavian Journal of Educational Research*, 51(5), 471-492.
- Sørli, M. A., & Ogden, T. (2015). School-wide positive behavior support–Norway: Impacts on problem behavior and classroom climate. *International Journal of School & Educational Psychology*, 3(3), 202-217.
- Sørli, M. A., Ogden, T., & Olseth, A. R. (2015). Preventing problem behavior in school through school-wide staff empowerment: Intervention outcomes. *World Journal of Educational Research*, 2(2), 117.
- Ward, B., & Gersten, R. (2013). A Randomized Evaluation of the Safe and Civil Schools Model for Positive Behavioral Interventions and Supports at Elementary Schools in a Large Urban School District. *School Psychology Review*, 42(3).
- Waschbusch, D. A., Pelham Jr, W. E., Massetti, G., & Northern Partners in Action for Children and Youth. (2005). The Behavior Education Support and Treatment (BEST) school intervention program: Pilot project data examining schoolwide, targeted-school, and targeted-home approaches. *Journal of Attention Disorders*, 9(1), 313-322.

#### General references cited in report

- Adolphus, K., Lawton, C. L., & Dye, L. (2013). The effects of breakfast on behavior and academic performance in children and adolescents. *Frontiers in human neuroscience*, 7, 425.
- Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, 12, 30-44. doi:<http://dx.doi.org/10.1016/j.edurev.2014.05.003>

- Ball, S. J., Maguire, M., & Braun, A. (2012). *How schools do policy: Policy enactments in secondary schools*. Routledge.
- Barmby, P. (2006) Improving teacher recruitment and retention: the importance of workload and pupil behaviour, *Educational Research*, 48(3), 247-265.
- Beaman, R., Wheldall, K., & Kemp, C. (2007). Recent research on troublesome classroom behaviour: A review. *Australasian Journal of Special Education*, 31(1), 45-60.
- Bennet, T. (2017). *Creating a Culture: How school leaders can optimise behaviour*, available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/602487/Tom\\_Bennett\\_Independent\\_Review\\_of\\_Behaviour\\_in\\_Schools.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/602487/Tom_Bennett_Independent_Review_of_Behaviour_in_Schools.pdf)
- Berg-Schlosser, D., & Cronqvist, L. (2005). Macro-quantitative vs. macro-qualitative methods in the social sciences—an example from empirical democratic theory employing new software. *Historical Social Research/Historische Sozialforschung*, 154-175.
- Bitsika, V. (2003). "But I'm Not Really Bad": Using an Ideographic Versus a Nomothetic Approach to Understand the Reasons for Difficult Behaviour in Children. *Journal of Psychologists and Counsellors in Schools*, 13(1), 87-98.
- Bodin, M. C., South, S. H., & Ingemarson, M. (2016). A Quasi-Randomized Trial of a School-Wide Universal Prevention Program: Results and Lessons Learned. *Scandinavian Journal of Educational Research*, 60(4), 449-476.
- Bosnjak, A., Boyle, C., & Chodkiewicz, A. R. (2017). An intervention to retrain attributions using CBT: A pilot study. *The Educational and Developmental Psychologist*, 34(1), 19-30.
- Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of school-wide positive behavioral interventions and supports on child behavior problems. *Pediatrics*, peds-2012.
- Brown, K. (2018). *Bullying: A review of the evidence*, available at: <https://epi.org.uk/publications-and-research/bullying-a-review-of-the-evidence/>
- Bruhn, A., McDaniel, S., & Kreigh, C. (2015). Self-monitoring interventions for students with behavior problems: A systematic review of current research. *Behavioral Disorders*, 40(2), 102-121.
- Cameron, R.J. (1998) School discipline in the United Kingdom: promoting classroom behaviour which encourages effective teaching and learning. *School Psychology Review*, 27(1), 33-44.
- Chodkiewicz, A. R., & Boyle, C. (2017). Positive psychology school-based interventions: A reflection on current success and future directions. *Review of Education*, 5(1), 60-86. doi: 10.1002/rev3.3080
- Chodkiewicz, A. R., & Boyle, C. (2016). Promoting positive learning in students aged 10 – 12 years using attribution retraining and cognitive behavioural therapy: A pilot study. *School Psychology International*, 37(5), 519-535. doi: 10.1177/0143034316667114
- Cooper, H., & Elton-Chalcraft, S. (Eds.). (2018). *Professional studies in primary education*. Sage.
- Daly-Smith, A. J., Zwolinsky, S., McKenna, J., Tomporowski, P. D., Defeyter, M. A., & Manley, A. (2018). Systematic review of acute physically active learning and classroom movement breaks on children's physical activity, cognition, academic performance and classroom behaviour: understanding critical design features. *BMJ open sport & exercise medicine*, 4(1), e000341.
- Day, C., Stobart, G., Sammons, P., Kington, A., Gu, Q., Smees, R., & Mujtaba, T. (2006). Variations in teachers' work, lives and effectiveness. *Final report for the VITAE Project*, DfES.
- Department for Education. (2019) *Pupil absence in schools in England: 2017 to 2018*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/787463/Absence\\_3term\\_201718\\_Text.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/787463/Absence_3term_201718_Text.pdf)
- Department for Education. (2018a) *A Review of School Exclusion: terms of reference*, available from: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/691005/A\\_Review\\_of\\_School\\_Exclusion-terms\\_of\\_reference.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/691005/A_Review_of_School_Exclusion-terms_of_reference.pdf)
- Department of Education (2018b) *Mental health and behaviour in schools*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/755135/Mental\\_health\\_and\\_behaviour\\_in\\_schools\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/755135/Mental_health_and_behaviour_in_schools_.pdf)
- Department for Education. (2016). *The link between absence and attainment at KS2 and KS4 2013/14 academic year*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/509679/The-link-between-absence-and-attainment-at-KS2-and-KS4-2013-to-2014-academic-year.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/509679/The-link-between-absence-and-attainment-at-KS2-and-KS4-2013-to-2014-academic-year.pdf)
- Department for Education. (2015) *School attendance parental responsibility measures Statutory guidance for local authorities, school leaders, school staff, governing bodies and the police*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/581539/School\\_attendance\\_parental\\_responsibility\\_measures\\_statutory\\_guidance.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/581539/School_attendance_parental_responsibility_measures_statutory_guidance.pdf)
- Department for Education (2012) *Pupil behaviour in schools in England*. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/184078/DFE-RR218.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/184078/DFE-RR218.pdf)

- Department for Education. (2010). *The importance of teaching: The schools white paper*. London: DfE.
- Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega Jr, R. (2014). Effects of the Good Behavior Game on challenging behaviors in school settings. *Review of educational research*, 84(4), 546-571.
- Garner, P. (2011). *Promoting the conditions for positive behaviour, to help every child succeed*. National College for School Leadership.
- Gastic, B. (2008). School truancy and the disciplinary problems of bullying victims. *Educational Review*, 60(4), 391-404.
- Graham-Matheson, L. (2014). *Essential Theory for Primary Teachers: An introduction for busy trainees*. Routledge.
- Glazzard, J., Denby, N., & Price, J. (2014). *Learning to Teach*. McGraw-Hill Education (UK).
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*, 336(7650), 924-926.
- Hershfeldt, P. A., Sechrest, R., Pell, K. L., Rosenberg, M. S., Bradshaw, C. P., & Leaf, P. J. (2009). Double-Check: A framework of cultural responsiveness applied to classroom behavior. *Teaching Exceptional Children Plus*, 6(2), 2-18.
- Higgins, J. P. T., Sterne, J. A. C., Savović, J., Page, M. J., Hróbjartsson, A., Boutron, I., ... & Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. *Cochrane Database Syst Rev*, 10(Suppl 1), 29-31.
- Jacobson, D. A. (2016). *Causes and effects of teacher burnout*. Walden University.
- Jennings, P. A., & Greenberg, M. T. (2009). The Prosocial Classroom: Teacher Social and Emotional Competence in Relation to Student and Classroom Outcomes. *Review of Educational Research*, 79(1), 491 - 525. doi:10.3102/0034654308325693
- Kokkinos, C. M. (2007) Job stressors, personality and burnout in primary school teachers, *British Journal of Educational Psychology*, 77(1), 229-243.
- Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (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(3), 643-680.
- Law, J., Plunkett, C. C., & Stringer, H. (2012). Communication interventions and their impact on behaviour in the young child: A systematic review. *Child Language Teaching and Therapy*, 28(1), 7-23.
- Lee, S. S. Y. (2014). Using fuzzy-set qualitative comparative analysis. *Epidemiology and health*, 36.
- Leithwood, K., & Day, C. (2008). *The impact of school leadership on pupil outcomes*. London: DCSF
- Loeber, R., & Farrington, D. P. (1999). *Serious and violent juvenile offenders: Risk factors and successful interventions*. Sage.
- McDermott, P. A., Mordell, M., & Stoltzfus, J. C. (2001). The organization of student performance in American schools: Discipline, motivation, verbal learning, nonverbal learning. *Journal of Educational Psychology*, 93(1), 65.
- McKenna, J. W., Flower, A., & Adamson, R. (2016). A systematic review of function-based replacement behavior interventions for students with and at risk for emotional and behavioral disorders. *Behavior modification*, 40(5), 678-712.
- Machalicek, W., O'Reilly, M. F., Beretvas, N., Sigafos, J., & Lancioni, G. E. (2007). A review of interventions to reduce challenging behavior in school settings for students with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 1(3), 229-246.
- Maggin, D. M., Chafouleas, S. M., Goddard, K. M., & Johnson, A. H. (2011). A systematic evaluation of token economies as a classroom management tool for students with challenging behavior. *Journal of School Psychology*, 49(5), 529-554.
- Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of school psychology*, 50(5), 625-654.
- Marzano, R., & Marzano, J. S. (2003). The key to classroom management. *Educational Leadership*, 61(1), 6-13.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, 151(4), 264-269.
- Moore, D. A., Russell, A. E., Matthews, J., Ford, T. J., Rogers, M., Ukoumunne, O. C., ... & Shaw, L. (2018). School-based interventions for attention-deficit/hyperactivity disorder: A systematic review with multiple synthesis methods. *Review of Education*, 6(3), 209-263.

- Nolan, K. M. (2011). Oppositional behavior in urban schooling: Toward a theory of resistance for new times. *International Journal of Qualitative Studies in Education*, 24(5), 559-572.
- Norwich, B., & Rovoli, I. (1993). Affective factors and learning behaviour in secondary school mathematics and English lessons for average and low attainers. *British Journal of Educational Psychology*, 63(2), 308-321.
- NSPCC (2018) *Our impact: How you're making a difference to children's lives*.  
[https://www.nspcc.org.uk/globalassets/blocks/about-us/our-organisation-and-structure/our-impact/nspcc\\_reporting-our-impact-2018.pdf](https://www.nspcc.org.uk/globalassets/blocks/about-us/our-organisation-and-structure/our-impact/nspcc_reporting-our-impact-2018.pdf)
- Núñez, J. L., & León, J. (2015). Autonomy support in the classroom. *European Psychologist*.
- Ofsted. (2014). *Below the Radar: Low Level Disruption in the Country's Classrooms*. Manchester: Ofsted.
- Ofsted (2019a) *State-funded schools inspections and outcomes as at 31 December 2018*  
<https://www.gov.uk/government/publications/state-funded-schools-inspections-and-outcomes-as-at-31-december-2018/state-funded-schools-inspections-and-outcomes-as-at-31-december-2018>
- Ofsted (2019b). *The education inspection framework*. Crown Copyright.
- Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive student behavior. *Society for Research on Educational Effectiveness*.
- Ozdemir, Y. (2007). The role of classroom management efficacy in predicting teacher burnout. *International Journal of Social Sciences*, 2(4), 257-263.
- Powell, S., & Tod, J. (2004). *A systematic review of how theories explain learning behaviour in school contexts*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Public Health England. (2015). Promoting children and young people's emotional health and wellbeing: A whole school and college approach.
- Rihoux, B. (2006). Qualitative comparative analysis (QCA) and related systematic comparative methods: Recent advances and remaining challenges for social science research. *International Sociology*, 21(5), 679-706.
- Rihoux, B., & Ragin, C. C. (2008). *Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques* (Vol. 51). Sage Publications.
- Sammons, P., Lindorff, A. M., Ortega, L., & Kington, A. (2016). Inspiring teaching: learning from exemplary practitioners. *Journal of Professional Capital and Community*, 1(2), 124-144.
- Scheuermann, B., & Hall, J. A. (2012). *Positive behavioral supports for the classroom*. Pearson/Merrill Prentice Hall.
- Smith, G. (2010). *Relation of Parental Social, Cognitive, and Behavioral Frameworks With School Behavior of High School Students*. Doctor of Education in Educational Leadership Thesis. University of Phoenix.
- Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Carpenter, J. R. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Bmj*, 355, i4919.
- Tanner-Smith, E. E., Durlak, J. A., & Marx, R. A. (2018). Empirically based mean effect size distributions for universal prevention programs targeting school-aged youth: A review of meta-analyses. *Prevention Science*, 19(8), 1091-1101.
- Thomas, J., O'Mara-Eves, A., & Brunton, G. (2014). Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. *Systematic reviews*, 3(1), 67.
- Timpson, E. (2019) *Timpson review of school exclusion*.  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/807862/Timpson\\_review.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/807862/Timpson_review.pdf)
- Ttofi, M., & Farrington, D. (2009). What works in preventing bullying: Effective elements of anti-bullying programmes. *Journal of Aggression, Conflict and Peace Research*, 1(1), 13-24.
- Watkins, C & Wagner, P (2000) *Improving school behaviour* London: Paul Chapman Publishing.
- Whear, R., Thompson-Coon, J., Boddy, K., Ford, T., Racey, D., & Stein, K. (2013). The effect of teacher-led interventions on social and emotional behaviour in primary school children: a systematic review. *British Educational Research Journal*, 39(2), 383-420.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary educational psychology*, 25(1), 68-81.
- Wilson, S. J., & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American journal of preventive medicine*, 33(2), S130-S143.
- Zyngier, D. (2007). Listening to teachers-listening to students: substantive conversations about resistance, empowerment and engagement. *Teachers and Teaching: theory and practice*, 13(4), 327-347.

## Appendix 1

Search Strategy used for Review 1

Database: PsycINFO <1806 to October Week 5 2018>

Search Strategy:

- 
- 1 (school\* or classroom\*).ti,ab. (403037)
  - 2 pupil\*.ti,ab. (25528)
  - 3 student\*.ti,ab. (473913)
  - 4 children.ti,ab. (464944)
  - 5 schoolchildren.ti,ab. (3309)
  - 6 2 or 3 or 4 or 5 (913264)
  - 7 ((behavior\* or behaviour\*) adj3 (school or class\*)).ti,ab. (17114)
  - 8 (behav\* adj (poor\* or bad\* or ill or challenging or aggressive or difficult)).ti,ab. (368)
  - 9 (willing\* adj2 learn\*).ti,ab. (315)
  - 10 ((ready or readiness) adj2 learn\*).ti,ab. (670)
  - 11 7 or 8 or 9 or 10 (18427)
  - 12 (model\* or framework\* or theor\* or reconstructivis\* or hypothes\*).ti. (234583)
  - 13 1 and 6 and 11 and 12 (390)

## Appendix 2. Quality appraisal for Review 1 studies

### Key

Quality rating	Criteria to guide assessment
LOW	Opinion, unsubstantiated arguments, restricted sample, eg one school.
MEDIUM	Small sample, eg two schools, appropriate analysis. Focus only on one variable's relationship with one measure of behaviour, literature review with unclear conclusions.
GOOD	Representative sample (national or large sample). Theory/framework/model tested. High quality review (although not systematic)
EXCELLENT	Systematic review. High quality.

**Table Quality appraisal for Review 1 - Aspects to Manage Directly**

Variable Name	Theme	Quality rating	Studies providing evidence
Teacher interest	Relationships	LOW	Daniels & Williams (2000), Liaisidou (2016) Núñez (2015)
	Relationships	MEDIUM	Bitsika (2003) Lyons (2006) Nolan (2011) Olvera (2008)
	Relationships	GOOD	Sullivan (2011) Hastings (2005)
Social Competency	Relationships	LOW	Mooij (1999)
	Relationships	MEDIUM	Eve (1978) Bear (1994) Kemp (2003) Maggio (2016) Richards (1989) Richards (1984)
	Relationships	GOOD	Pingault (2015)
Relationship with Peers	Relationships	LOW	Cornwall (2015) Daniels & Williams (2000)
	Relationships	MEDIUM	Bear & Rys (1994) Bidell (2010) LaRoque (2008) Lopez (2005) Svendham (1994)
	Relationships	GOOD	Cheung (1997) Chirinos (2018) Sullivan (2011)
Relationship Development	Relationships	LOW	Daniels & Williams (2000) Liaisidou (2016)
	Relationships	MEDIUM	Haynes (1990) Jackson (1998) Lyons (2006) Svendham (1994)
	Relationships	GOOD	Heaven (2009) Sullivan (2011)
Teacher Connection	Relationships	LOW	Cornwall (2015) Hershfeldt (2009) Núñez (2015)



Variable Name	Theme	Quality rating	Studies providing evidence
	Relationships	MEDIUM	Bitsika (2003) Haynes (1990) Hastings (2005) Nolan (2011)
	Relationships	GOOD	Heaven (2009) Nie (2009)
Academic Achievement	Teaching and Learning	MEDIUM	Cadieux (2003) Kerr (2000)
	Teaching and Learning	GOOD	Adolphus (2013) Boon (2011) Chirinos (2018) McDermott (2001)
Learning Potential	Teaching and Learning	LOW	Mooij (1999)
	Teaching and Learning	GOOD	Garner (2011)
Educational opportunity	<b>Teaching and Learning</b>	MEDIUM	Eve (1978)
	<b>Teaching and Learning</b>	GOOD	Dever (2016) Sullivan (2011)
Connection to Curriculum	<b>Teaching and Learning</b>	LOW	Cornwall (2015) Ellis & Tod (2018) Hershfeldt (2009)
	<b>Teaching and Learning</b>	MEDIUM	Borders (2004) Gottfredson (1994).

Variable Name	Theme	Quality rating	Studies providing evidence
Exclusions/Suspensions	<b>School organisation</b>	LOW	Liaisidou (2016) Smith (2010)
School Environment	<b>School organisation</b>	LOW	Cornwall (2015) Daniels & Williams (2000) Liaisidou (2016) Mooij (1999)
	<b>School organisation</b>	MEDIUM	Bitsika (2003) Lyons (2006) Nolan (2011) Tremblay (2010)
	<b>School organisation</b>	GOOD	Smith (2010) Sullivan (2011)
	<b>School organisation</b>	EXCELLENT	Finn (2003)
School Ethos/Policies	<b>School organisation</b>	LOW	Cornwall (2015) Daniels & Williams (2000) Hershfeldt (2009) Ellis & Tod (2018) Liaisidou (2016)
	<b>School organisation</b>	MEDIUM	Cadieux (2003) Nolan (2011) Norwich (1993)
	<b>School organisation</b>	GOOD	Garner (2011) Smith (2010)
School Transitions	<b>School organisation</b>	LOW	Ellis & Tod (2018)
School leadership	<b>School organisation</b>	LOW	Daniels & Williams (2000)
	<b>School organisation</b>	GOOD	Garner (2011)
Behaviour management approach	<b>School organisation</b>	LOW	Daniels & Williams (2000) Hershfeldt (2009) Liaisidou (2016) Lyons (2006) Núñez (2015)

Variable Name	Theme	Quality rating	Studies providing evidence
	<b>School organisation</b>	MEDIUM	Bitsika (2003) Cadieux (2003) DiStefano (2008) Hastings (2005) Nie (2009) Norwich (1993)
	<b>School organisation</b>	GOOD	Garner (2011) Gottfredson (1994) Smith (2010)
Definition of appropriate behaviour	<b>School organisation</b>	MEDIUM	Bitsika (2003) Hastings (2005)
School & community relations	<b>School organisation</b>	LOW	Cornwall (2015) Daniels & Williams (2000) Hershfeldt (2009)
	<b>School organisation</b>	MEDIUM	Hastings (2005) Nolan (2011)
	<b>School organisation</b>	GOOD	Garner (2011) Smith (2010) Sullivan (2011)
Class Size	<b>School organisation</b>	EXCELLENT	Finn (2003)
Educational Motivation	<b>Attitudes and self-concept</b>	LOW	Cornwall (2015) Núñez (2015)
	<b>Attitudes and self-concept</b>	MEDIUM	Bidell (2010)
	<b>Attitudes and self-concept</b>	GOOD	Adolphus (2013) Dever (2016)

Variable Name	Theme	Quality rating	Studies providing evidence
Academic Expectations	Attitudes and self-concept	MEDIUM	Borders (2004) Norwich (1993)
	Attitudes and self-concept	GOOD	Chirinos (2018) Dever (2016)
Attitudes to Learning	Attitudes and self-concept	LOW	Cornwall (2015) Daniels & Williams (2000)
	Attitudes and self-concept	MEDIUM	Borders (2004) Kemp (2003) Norwich (1993)
	Attitudes and self-concept	GOOD	Dever (2016) Nie (2009) Sullivan (2011)
Labelling	Attitudes and self-concept	GOOD	Cheung (1997)
Happiness	Attitudes and self-concept	MEDIUM	Bidell (2010) Kerr (2000)
Personal Perception of Consequences	Attitudes and self-concept	LOW	Cornwall (2015)
	Attitudes and self-concept	MEDIUM	LaRoque (2008)
	Attitudes and self-concept	GOOD	Sullivan (2011)
Self-evaluation	Attitudes and self-concept	LOW	Cornwall (2015) Daniels & Williams (2000) Hershfeldt (2009)

Variable Name	Theme	Quality rating	Studies providing evidence
	Attitudes and self-concept	MEDIUM	Bidell (2010) Eve (1978) Haynes (1990) Jackson (1998) Lopez (2005)
Emotional regulation	Attitudes and self-concept	LOW	Daniels & Williams (1999) Ellis & Tod (2018) Mooij (1999)
	Attitudes and self-concept	MEDIUM	Bidell (2010) Jackson (1998) Lochman (2002) Lopez (2005) Lyons (2006) Maggio (2016) Nelson (1999)
Attributions	Attitudes and self-concept	LOW	Liaisidou (2016) Núñez (2015)
	Attitudes and self-concept	MEDIUM	Bitsika (2003) Eve (1978) Hastings (2005) Hershfeldt (2009) Lochman (2002)
	Attitudes and self-concept	GOOD	Chirinos (2018) Dever (2016) Heaven (2009)

**Table Quality appraisal for Review 1 - Aspects to Identify and Influence**

Variable Name	Theme	Quality rating	Studies providing evidence
Choice of peers	Relationships out of school	MEDIUM	Svendham (1994) LaRoque (2008) Lopez (2005)
	Relationships out of school	GOOD	Sullivan (2011) Cheung (1997)
Relationship with Others	Relationships out of school	LOW	Daniels & Williams (1999) Ellis & Tod (2018)
	Relationships out of school	MEDIUM	Bear & Rys (1994) Jackson (1998) Kemp (2003) Lochman (2002)
	Relationships out of school	GOOD	Sullivan (2011)
Peer Group Perception of Consequences	Relationships out of school	MEDIUM	LaRoque (2008)
Family relationships	Relationships out of school	LOW	Ellis & Tod (2018)
	Relationships out of school	MEDIUM	Haynes (1990) Kemp (2003) Güleç (2011)
	Relationships out of school	GOOD	Caughy (2007) Lorber (2011) Sullivan (2011)
Social Support	Relationships out of school	MEDIUM	Jackson (1998)
	Relationships out of school	GOOD	Knutson (2004) Liao (2015)
Discipline	Home life	MEDIUM	Lochman (2002)
	Home life	GOOD	Knutson (2004) Kemp (2003)
Parental Involvement in Homework	Home life	MEDIUM	Bitsika (2003) Lochman (2002)
	Home life	GOOD	Caughy (2007) Smith (2010)

<b>Variable Name</b>	<b>Theme</b>	<b>Quality rating</b>	<b>Studies providing evidence</b>
Negative Life-Events	Stressors	MEDIUM	Jackson (1998) Kerr (2000)
	Stressors	GOOD	Boon (2011) Liao (2015) Lorber (2011)
Response	Stressors	MEDIUM	Eve (1978) Maggio (2016)
	Stressors	GOOD	Boon (2011)
Adaptive Functioning	Stressors	MEDIUM	DiStefano (2008) Eve (1978) Jackson (1998) Maggio (2016) Olvera (2008)
	Stressors	GOOD	Boon (2011) Heaven (2009) Kerr (2000) McDermott (2001)
Emotional Functioning	Stressors	MEDIUM	Jackson (1998) Olvera (2008)
	Stressors	GOOD	Heaven (2009)
Behavioural Functioning	Stressors	MEDIUM	Eve (1978) Jackson (1998) Lyons (2006) Olvera (2008)

<b>Variable Name</b>	<b>Theme</b>	<b>Quality rating</b>	<b>Studies providing evidence</b>
Behavioural Functioning	Stressors	GOOD	Heaven (2009) Kerr (2000) Pingault (2015)
Coping and resilience	Stressors	MEDIUM	DiStefano (2008) Gottfredson (1994). Lochman (2002)
	Stressors	GOOD	Boon (2011)
Anti-Social Behaviour	Behaviour out of school	LOW	Mooij (1999)
	Behaviour out of school	GOOD	Heaven (2009)
Incarceration	Behaviour out of school	LOW	Slee (2014)
Intellectual Ability	Ability	LOW	Mooij (1999)
	Ability	MEDIUM	Borders (2004) Jackson (1998) Kerr (2000)
Breakfast	Nutrition	EXCELLENT	Adolphus (2013)
Social and welfare services	Out of school support	LOW	Daniels & Williams (1999) Ellis & Tod (2018)
Social and welfare services	Out of school support	GOOD	Pingault (2015)

**Table Quality appraisal for Review 1 - Aspects to be Aware of.**

<b>Variable Name</b>	<b>Theme</b>	<b>Quality rating</b>	<b>Studies providing evidence</b>
Sociometric Status	Home life	LOW	Cornwall (2015) Hershfeldt (2009) Mooij (1999)
	Home life	MEDIUM	Jackson (1998) Richards (1989) Richards (1984) Tremblay (2010)
	Home life	GOOD	Bear & Rys (1994) Caughy (2007) Knutson (2004) Sullivan (2011)
Family Functioning	Home life	LOW	Cornwall (2015) Hershfeldt (2009) Ellis & Tod (2018) Mooij (1999)
	Home life	MEDIUM	Dinh (2001) Güleç (2011) Kemp (2003) Nelson (1999) Olvera (2008) Tremblay (2010)
	Home life	GOOD	Boon (2011) Caughy (2007) Cheung (1997) Lorber (2011)
Witnessing Violence	Home life	MEDIUM	Güleç (2011)
	Home life	GOOD	Liao (2015)
Parental Mental Health and Wellbeing	Home life	GOOD	Caughy (2007)
Parental Education Level	Home life	GOOD	Boon (2011) Caughy (2007)
Parental view of Education	Home life	MEDIUM	Güleç (2011)
	Home life	GOOD	Chirinos (2018)
Abuse	Home life	GOOD	Knutson (2004)
Culture Conflict	Culture	LOW	Cornwall (2015) Ellis & Tod (2018) Mooij (1999)
	Culture	MEDIUM	Bitsika (2003) Eve (1978) Richards (1989)
Acculturation	Culture	MEDIUM	Bear & Rys (1994) Dinh (2001) Hershfeldt (2009)

### Appendix 3. Search strategy used for Reviews 2&3

Database: PsycINFO <1806 to September Week 4 2018>

Search Strategy:

- 
- 1 school\*.ti,ab. (359749)
  - 2 pupil\*.ti,ab. (25485)
  - 3 student\*.ti,ab. (472333)
  - 4 children.ti,ab. (463491)
  - 5 schoolchildren.ti,ab. (3299)
  - 6 2 or 3 or 4 or 5 (910328)
  - 7 classroom\*.ti,ab. (80486)
  - 8 school based.ti,ab. (13632)
  - 9 whole school.ti,ab. (912)
  - 10 7 or 8 or 9 (93179)
  - 11 teacher\*.ti,ab. (170881)
  - 12 (manag\* adj4 behavio?r\*).ti,ab. (13959)
  - 13 (strateg\* adj5 (manag\* or control\* or improv\* or modif\*)).ti,ab. (35017)
  - 14 intervention\*.ti,ab. (337801)
  - 15 (approach or approaches).ti,ab. (470734)
  - 16 11 or 12 or 13 or 14 or 15 (909454)
  - 17 (behavior\* or behaviour\*).ti,ab. (870629)
  - 18 comparison.ti. (53006)
  - 19 (comparison adj2 (children or group\*)).ab. (16158)
  - 20 control group.ab. (63698)
  - 21 experiment\*.ab. (378749)
  - 22 random\*.ab. (178384)
  - 23 systematic\*.ti,ab. (109351)
  - 24 effectiveness.ti,ab. (140375)
  - 25 18 or 19 or 20 or 21 or 22 or 23 or 24 (802914)
  - 26 1 and 6 and 10 and 16 and 17 and 25 (2875)

Appendix 4 Data table for QCA

Study/case	Targets misbehaviour	Targets prosocial behaviour	Targets learning behaviour	Tackles academic issues	Teaches coping and resilience	Tailored to individuals	Improves relationships	Whole school	Home/School communication	Includes reward system	Intensive teacher training	Teachers consult with experts
Kamps 2015	0.33	1	0	0	0.33	0	0.33	0.33	0	1	0.33	0
Leflot 2010	0.33	1	0	0.33	0	0	0.33	0.33	0	1	0.33	0.66
Humphrey 2018	0.33	1	0.33	0.33	1	0	1	1	1	0	0.67	0
Aasheim 2018	0.66	0.66	1	0	1	0	1	0.33	0	0	1	0
Evertson (1989)	0.33	0.33	1	1	0.66	0.33	0.33	0.33	0	1	0.33	1
Fernandez (2015)	1	1	0.33	0	0	1	0.33	0	0	1	0.33	0.66
Ford (2018)	0.66	0.66	1	0	1	0	1	0.66	0	0	1	0
Fossum (2017)	0.66	0.66	1	0	1	0	1	0.66	0	0	1	0
Han (2005)	0.33	1	0	0.33	1	0	1	0.66	1	1	1	1
Hickey (2017)	0.66	0.66	1	0	1	0	1	0.66	0	0	1	0
Hutchings (2013)	0.66	0.66	1	0	1	0	1	0.66	0	0	1	0
McGilloway (2010)	0.66	0.66	1	0	1	0	1	0.66	0	0	1	0
Murray (2018)	0.66	0.66	1	0.33	1	0	1	0.66	0	0	1	0
Piwovar (2013)	0.33	0.33	0.33	0.33	0.33	0.66	0.66	0.33	0	0.33	0.67	0
Capella (2012)	1	0.66	1	1	1	0.33	1	0.33	0	0	0	1
Bodin (2016)	0.33	1	0.33	0	0	0	0.33	1	1	1	0.33	0
Gottfredson (1993)	0.66	1	0.33	0.33	0	0.33	0	1	1	1	0.55	0
Nelson (1996)	1	1	0.33	0.33	1	0.33	1	1	1	0	0.55	0.66
Sorlie (2007)	0.66	1	0	1	1	0.33	0.33	1	1	0.33	0.55	0.66
Waschbush (2005)	1	1	1	1	1	1	1	1	1	1	0.55	0.33

<b>Condition</b>	<b>Coding</b>	<b>Criteria</b>
<b>Targets misbehaviour</b>	0	No focus on misbehaviour
	0.33	Focusses on misbehaviour e.g. by improving or modelling good behaviour
	0.66	Focusses on tackling misbehaviour - to an extent identifies what are problem behaviours and seeks to tackle them either by discipline or training behaviours.
	1	Focusses on tackling misbehaviour - clearly identifies problem behaviours and successfully tackles them either by discipline or training behaviours.
<b>Targets prosocial behaviour</b>	0	Not at all
	0.33	Intervention probably benefits prosocial behaviour
	0.66	Intervention has an indirect benefit to prosocial behaviour or social behaviour that can be regarded as prosocial behaviour.
	1	Intervention successfully targets prosocial behaviour or social behaviour that can be regarded as prosocial behaviour.
<b>Targets learning behaviour</b>	0	No focus on learning behaviour
	0.33	Probably improves good learning behaviour with unclear outcomes.
	0.66	Some focus on learning behaviours which improves them
	1	Clear focus on addressing learning behaviours - identifies what are inappropriate learning behaviours and successfully develops new learning behaviours.
<b>Tackles academic issues</b>	0	No discernible influence on academic issues
	0.33	An indirect focus on academic issues but no sign they improve or some in some classes which do improve
	0.66	An indirect focus which improves whole school academic issues or a direct focus on some classes which do improve
	1	Intervention targets and improves whole school academic issues
<b>Teaches coping and resilience</b>	0	No indirect or direct focus
	0.33	Intervention indirectly benefits some children's coping and or resilience
	0.66	Some focus on coping and resilience
	1	A clear focus on coping and resilience which is stated as benefitting children
<b>Tailored to individuals</b>	0	Everyone is treated the same - universal
	0.33	A universal intervention with some flexibility between classes/individuals
	0.66	The intervention is only for certain groups of students, or within a universal or targeted intervention certain students get an intervention tailored to their needs
	1	All participants receive an intervention that is tailored to their needs in some way
<b>Improves relationships</b>	0	No relationships in school are likely to be improved by the intervention
	0.33	There is likely to some improvement in relationships but it is not targeted.
	0.66	It is clear that indirectly the intervention has some benefit to relationships
	1	At least some of intervention directly aims to improve relationships
<b>Whole school</b>	0	Could not be focussed at whole school
	0.33	Could be extended to whole school but focussed on individual classes
	0.66	Could be extended to whole school but focussed on year groups
	1	Focussed at whole school
	0	No evidence of school/home interaction (other than child participation consent)



<b>Home/School communication</b>	0.33	The maybe some school / home interaction e.g. survey
	0.66	Clear indirect home school effects
	1	Addresses school / home interaction
<b>Includes reward system</b>	0	Complete absence of reward system
	0.33	No direct reward process but child likely to be rewarded
	0.66	Intervention includes reward system but unclear extent of take-up
	1	Reward system tested that improves behaviours
<b>Intensive teacher training</b>	0	Less than 11 hours of teacher training
	0.33	11-20 hours of teacher training
	0.66	21-44 hours of teacher training
	1	45+ hours of teacher training
<b>Teachers consult with experts</b>	0	Not at all
	0.33	Some teachers may have consulted with experts beyond the training process
	0.66	Most teachers probably consulted with experts beyond the training process.
	1	All teacher consulted with experts beyond the training process