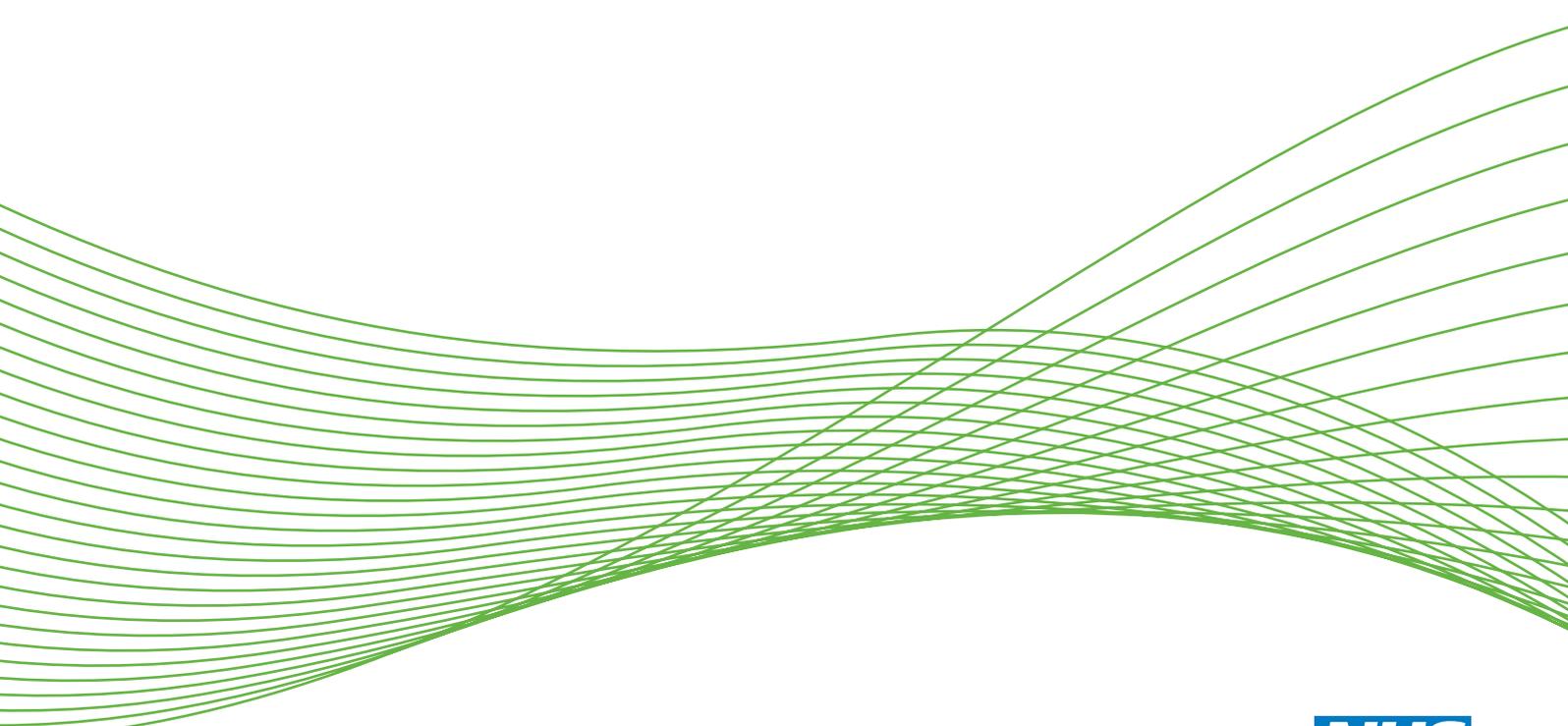


Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis

C Bonell, F Jamal, A Harden, H Wells, W Parry, A Fletcher, M Petticrew, J Thomas, M Whitehead, R Campbell, S Murphy and L Moore



**National Institute for
Health Research**

Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis

C Bonell,^{1*} F Jamal,² A Harden,² H Wells,³ W Parry,⁴
A Fletcher,⁵ M Petticrew,³ J Thomas,⁶ M Whitehead,⁷
R Campbell,⁸ S Murphy⁵ and L Moore⁵

¹Centre for Evidence Based Intervention, Department of Social Policy and Intervention, University of Oxford, Oxford, UK

²Institute for Health and Human Development, University of East London, London, UK

³Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK

⁴Department of Quantitative Social Science, Institute of Education, University of London, London, UK

⁵DECIPHER, School of Social Sciences, Cardiff University, Cardiff, UK

⁶Department of Childhood, Families and Health, Institute of Education, University of London, London, UK

⁷Institute of Psychology, Health and Society, University of Liverpool, Liverpool, UK

⁸School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author

Declared competing interests of authors: Adam Fletcher, James Thomas and Margaret Whitehead have received grants to their institution from the National Institute for Health Research (NIHR). Rona Campbell has received grants to her institution, support for travel and payment for writing/reviewing from the NIHR Public Health Research programme.

Disclaimer: This report contains transcripts from interviews synthesised in the course of the research which may contain language that may offend some readers.

Published June 2013

DOI: 10.3310/phr01010

This report should be referenced as follows:

Bonell C, Jamal F, Harden A, Wells H, Parry W, Fletcher A, *et al.* Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Res* 2013;**1**(1).

Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full PHR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/phr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Public Health Research* journal

Reports are published in Public Health Research (PHR) if (1) they have resulted from work for the PHR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in Public Health Research are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

PHR programme

The Public Health Research (PHR) programme, part of the National Institute for Health Research (NIHR), evaluates public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health. The Public Health Research programme also complements the NIHR Health Technology Assessment programme which has a growing portfolio evaluating NHS public health interventions.

For more information about the PHR programme please visit the website: www.phr.nihr.ac.uk/

This report

The research reported in this issue of the journal was funded by the PHR programme as project number 09/3002/08. The contractual start date was in March 2010. The final report began editorial review in March 2012 and was accepted for publication in October 2012. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PHR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the PHR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2013. This work was produced by Bonell *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Public Health Research Editor-in-Chief

Professor Catherine Law Professor of Public Health and Epidemiology, Unit Head, Centre for Paediatric Epidemiology and Biostatistics, UCL Institute of Child Health, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Health Sciences, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Tom Marshall Reader in Primary Care, School of Health and Population Sciences, University of Birmingham, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Honorary Professor, Business School, Winchester University and Medical School, University of Warwick, UK

Professor Jane Norman Professor of Maternal and Fetal Health, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, NICE, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professorial Research Associate, University College London, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis

C Bonell,^{1*} F Jamal,² A Harden,² H Wells,³ W Parry,⁴ A Fletcher,⁵ M Petticrew,³ J Thomas,⁶ M Whitehead,⁷ R Campbell,⁸ S Murphy⁵ and L Moore⁵

¹Centre for Evidence Based Intervention, Department of Social Policy and Intervention, University of Oxford, Oxford, UK

²Institute for Health and Human Development, University of East London, London, UK

³Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK

⁴Department of Quantitative Social Science, Institute of Education, University of London, London, UK

⁵DECIPHer, School of Social Sciences, Cardiff University, Cardiff, UK

⁶Department of Childhood, Families and Health, Institute of Education, University of London, London, UK

⁷Institute of Psychology, Health and Society, University of Liverpool, Liverpool, UK

⁸School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author

Background: In contrast to curriculum-based health education interventions in schools, the school environment approach promotes health by modifying schools' physical/social environment. This systematic review reports on the health effects of the school environment and processes by which these might occur. It includes theories, intervention outcome and process evaluations, quantitative studies and qualitative studies.

Research questions: Research question (RQ)1: What theories are used to inform school environment interventions or explain school-level health influences? What testable hypotheses are suggested? RQ2: What are the effects on student health/inequalities of school environment interventions addressing organisation/management; teaching/pastoral care/discipline; and the physical environment? What are the costs? RQ3: How feasible/acceptable and context dependent are such interventions? RQ4: What are the effects on student health/inequalities of school-level measures of organisation/management; teaching/pastoral care/discipline; and the physical environment? RQ5: Through what processes might such influences occur?

Data sources: A total of 16 databases were searched between 30 July 2010 and 23 September 2010 to identify relevant studies, including the British Educational Index, the Cumulative Index to Nursing and Allied Health Literature, the Health Management Information Consortium, EMBASE, MEDLINE and PsycINFO. In addition, references of included studies were checked and authors contacted.

Review methods: In stage 1, we mapped references concerning how the school environment affects health and consulted stakeholders to identify stage 2 priorities. In stage 2, we undertook five reviews corresponding to our RQs.

Results: Stage 1: A total of 82,775 references were retrieved and 1144 were descriptively mapped. Stage 2: A total of 24 theories were identified (RQ1). The human functioning and school organisation, social capital and social development theories were judged most useful. Ten outcome evaluations were included (RQ2). Four US randomised controlled trials (RCTs) and one UK quasi-experimental study examined interventions building school community/relationships. Studies reported benefits for some, but not all outcomes (e.g. aggression, conflict resolution, emotional health). Two US RCTs assessed interventions empowering students to contribute to modifying food/physical activity environments, reporting benefits for physical activity but not for diet. Three UK quasi-experimental evaluations examined playground improvements, reporting mixed findings, with benefits being greater for younger children and longer break times. Six process evaluations (RQ3) reported positively. One study suggested that implementation was facilitated when this built on existing ethos and when senior staff were supportive. We reviewed 42 multilevel studies, confining narrative synthesis to 10 that appropriately adjusted for confounders. Four UK/US reports suggested that schools with higher value-added attainment/attendance had lower rates of substance use and fighting. Three reports from different countries examined school policies on smoking/alcohol, with mixed results. One US study found that schools with more unobservable/unsupervised places reported increased substance use. Another US study reported that school size, age structure and staffing ratio did not correlate with student drinking. Twenty-one qualitative reports from different countries (RQ5) suggested that disengagement, lack of safety and lack of participation in decisions may predispose students to engage in health risks.

Limitations: We found no evidence regarding health inequalities or cost, and could not undertake meta-analysis.

Conclusions: There is non-definitive evidence for the feasibility and effectiveness of school environment interventions involving community/relationship building, empowering student participation in modifying schools' food/physical activity environments, and playground improvements. Multilevel studies suggest that schools that add value educationally may promote student health. Qualitative studies suggest pathways underlying these effects. This evidence lends broad support to theories of social development, social capital and human functioning and school organisation. Further trials to examine the effects of school environment modifications on student health are recommended.

Funding: The National Institute for Health Research Public Health Research programme.

Contents

List of abbreviations	ix
Executive summary	xi
Chapter 1 Background	1
The importance of the school environment	1
Rationale for the systematic review	2
Chapter 2 Aim and research objectives	5
Research question 1	5
Research question 2	5
Research question 3	5
Research question 4	5
Research question 5	6
Protocol	6
Chapter 3 Report organisation and terminology	7
Chapter 4 Data management	9
Stage 1 Identifying and describing the references	11
Chapter 5 Evidence map, theory map and stakeholder consultations	13
Aim and research questions	13
Methods	13
Results	17
Implications for stage 2: in-depth review	23
Stage 2 In-depth synthesis	25
Chapter 6 Research question 1: theory synthesis	27
Research question	27
Methods	27
Overview of included reports	28
Results	29
Discussion	32
Chapter 7 Research question 2: outcome evaluations	35
Research question	35
Methods	35
Overview of included reports	37
Results	41
Discussion	44

Chapter 8 Research question 3: process evaluations	47
Research question	47
Methods	47
Overview of included reports	48
Results	49
Discussion	50
Chapter 9 Research question 4: multilevel studies	51
Research question	51
Methods	51
Overview of included reports	54
Results	58
Discussion	61
Chapter 10 Research question 5: qualitative studies	65
Research question	65
Methods	65
Overview of included reports	67
Results	69
Discussion	80
Chapter 11 Discussion and conclusions	83
Review limitations	83
Key results	85
Overall synthesis	90
Conclusions	96
Implications	97
Acknowledgments	103
References	105
Appendix 1 Searches	117
Appendix 2 Included studies/theory for stage 2 in-depth review	133
Appendix 3 Stage 1 evidence and theory map countries listed in included references	139
Appendix 4 Stage 1 theory map	141
Appendix 5 Data extraction and quality appraisal tools	147
Appendix 6 Data extraction and quality appraisal tables	173
Appendix 7 Protocol	295

List of abbreviations

AAYP	Aban Aya Youth Project (intervention)	OR	odds ratio
ALPHA	Advice Leading to Public Health Advancement	OSDUS	Ontario Student Drug Use Survey
ASB	antisocial behaviour	PE	physical education
BMI	body mass index	RCT	randomised controlled trial
CAPSLE	Creating a Peaceful School Learning Environment (intervention)	RQ	research question
CAS	Control of Adolescent Smoking	SCI	school/community intervention
CCTV	closed-circuit television	SD	standard deviation
CDP	Child Development Project (intervention)	SDC	social development curriculum
CI	confidence interval	SES	socioeconomic status
DECIPHer	Development and Evaluation of Complex Interventions for Public Health Improvement	SHAPES	School Health Action Planning and Evaluation System
GCSE	General Certificate of Secondary Education	VPA	vigorous physical activity
HPS	Health Promoting Schools	WHO	World Health Organization
HSE	Healthy School Ethos (intervention)		
HYP	Healthy Youth Places (intervention)		
ICC	intraclass correlation coefficient		
M-SPAN	Middle-School Physical Activity and Nutrition (intervention)		
MVPA	moderate/vigorous physical activity		
NELS	National Education Longitudinal Study		
NLSAH	National Longitudinal Study of Adolescent Health		

Researchers

AF	Adam Fletcher
AH	Angela Harden
CB	Chris Bonell
CV	Carol Vigurs
FJ	Farah Jamal
HW	Helene Wells
MP	Mark Petticrew
RL	Rebecca Langford
WP	Will Parry

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background and rationale

The health of young people in the UK is among the worst in Europe. The effects of curriculum-based health education interventions in schools show mixed results. A complementary 'school environment' approach has been used instead to modify the school physical and social and cultural environment to promote health. This report presents a systematic review of school environment studies addressing multiple questions and using diverse types of evidence.

School environment interventions are supported by the World Health Organization (WHO) framework for Health Promoting Schools (HPS). A Cochrane review of HPS interventions (which address school environment alongside parent/community involvement and curriculum) is under way. Our review is different because it focuses on interventions addressing the school environment alone in order to isolate environmental effects (which is not possible when combining environment with curriculum components). Process evaluation studies are also reviewed, as they are useful for informing decisions about the wider implementation of interventions. Because health outcomes also vary between schools in the absence of specific interventions, and research suggests that these differences are attributable to school-level measures of the school social and physical environment, we have also included quantitative studies of school-level effects in our review. Although existing reviews have examined such research, they have not drawn authoritative conclusions because of methodological limitations in the studies they have included. We have therefore applied more rigorous inclusion criteria to review quantitative studies of school-level health effects. We also review qualitative studies examining the processes underlying such effects.

Aim and research questions

This systematic review aims to synthesise evidence relating to the health effects of school environment interventions and of school-level measures of the social and physical environment and the processes underlying these. The review was conducted in two stages. In stage 1, we identified and descriptively mapped a broad array of potentially relevant literature, including research involving all aspects of the school environment and student health as well as teacher health. Stage 2 focused specifically on student health and defined the school environment more narrowly in terms of how schools are organised/managed, how they teach, how they provide pastoral care and discipline and/or the school physical environment. It involved five in-depth reviews to address the following research questions (RQs):

Research question 1

What theories and conceptual frameworks are most commonly used to inform school environment interventions or explain school-level influences on health? What testable review hypotheses do these suggest?

Research question 2

What are the effects on health and health inequalities among school students aged 4–18 years of school environment interventions (modifying how schools are organised/managed, how they teach, provide pastoral care to and discipline students, and/or the school physical environment) that do not include health education or health services as intervention components and which are evaluated using prospective experimental and quasi-experimental designs? What are their direct and indirect costs?

Research question 3

How feasible and acceptable are the school environment interventions examined in studies addressing RQ2? How does context affect this?

Research question 4

What are the effects on health and health inequalities among school students aged 4–18 years of school-level measures of school organisation/management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers, and/or the physical environment (measured using 'objective' data rather than aggregate self-reports from the same individuals who provide data on outcomes), examined using multilevel quantitative designs?

Research question 5

Through what processes might these school-level influences occur, examined using qualitative research?

We review each of the five RQs in separate chapters. We then assess the review hypotheses developed under RQ1 in relation to the totality of empirical evidence in our final chapter's overall synthesis.

Methods**Stage 1: identifying and describing the literature**

To locate evidence and theory, 16 databases were searched between 30 July 2010 and 23 September 2010, including the British Educational Index, the Cumulative Index to Nursing and Allied Health Literature, the Health Management Information Consortium, EMBASE, MEDLINE and PsycINFO. A priori criteria were developed to identify relevant references based on title and abstract, and these were descriptively coded (e.g. country of study, health topic, school level) to develop an evidence and theory map. We consulted with key stakeholders, including young people, about the map and the implications for stage 2.

Stage 2: in-depth synthesis

An in-depth synthesis was conducted for each of the five RQs. Specific exclusion criteria, quality assessment and data extraction tools were developed for each synthesis. Additional searches were conducted by checking references of included reports and contacting study authors. A narrative synthesis approach was used for RQ1–4 and a meta-ethnography approach was used for RQ5.

Results

A total of 1144 references were included in the evidence and theory map. Most were references to primary research conducted in high-income countries. The main health topics identified at the mapping stage were student violence, bullying, harassment, diet and physical activity. The main aspects of the school environment identified were school management/policies, catering services/vending machines and sport/active transport.

The findings of the theory map and the consultations with key stakeholders suggested that the most important school environment interventions and determinants to focus on were those relating to how schools are organised and managed, how they deliver teaching, pastoral care and discipline, and schools' physical environments.

Research question 1: theory synthesis

A total of 24 theories were cited in either stand-alone theory papers or empirical reports addressing RQ2–5. The most commonly cited theories were ecological systems theory (cited in 10 reports), social control theory ($n=6$), social disorganisation theory ($n=5$), social learning theory ($n=5$), theory of human functioning and school organisation ($n=5$) and social cognitive theory ($n=4$).

Inclusion criteria were developed to assess which theories should inform our primary and secondary review hypotheses. Three theories informed our primary review hypotheses:

- social capital theories – schools will foster health by having a stable student and staff body, good relationships between staff and students and a positive school ethos of stable shared norms
- social development model – schools reduce antisocial behaviour by providing opportunities for students to participate fully in learning and community life, develop the skills necessary for such participation and ultimately gain recognition
- theory of human functioning and school organisation – schools foster student autonomy and health by reducing social boundaries between staff and students and among students, and ensuring student-centred framing of learning, management and other school systems.

Research question 2: outcome evaluations

A total of 16 reports of 10 studies were included that evaluated the outcomes of interventions aiming to modify the school environment without simultaneously addressing school health curricula. Of these 10 studies, six were randomised controlled trials (RCTs) and four were quasi-experimental studies. Across all reports, more measures were reported as providing significant benefits than as not significantly affecting outcomes, and none reported significant harms.

Five outcome evaluations examined interventions that encouraged staff and students to build a stronger sense of community and/or better interpersonal relations at school. Such studies have been conducted in a range of school settings in elementary, middle and secondary/high schools. All except the Healthy School Ethos (HSE) intervention (UK) were conducted in the USA. Evaluations reported benefits regarding some but not all measures of emotional health, conflict resolution, aggression, victimisation and perceived student safety. However, the strongest evaluation in this category, the Aban Aya Youth Project (AAYP), found school environment change to be associated with fewer significant health benefits than curriculum only.

Two RCTs assessed an intervention that combined changes to US middle schools' food and physical activity environments alongside actions which aimed to empower students to contribute to achieving these changes. These were well conducted and both reported intervention benefits for student physical activity but not for healthy eating. The mediation analysis in the Healthy Youth Places (HYP) study suggested that student empowerment partly explained intervention effects. Three quasi-experimental evaluations of an intervention to improve playgrounds in British primary schools reported mixed findings on students' physical activity with indications that benefits were greater for younger children and when break time was longer.

The outcome evaluation studies provide little information on the likely impact of school environment interventions on health inequalities. Two studies of playground interventions reported costs although none reported on cost-effectiveness.

Research question 3: process evaluations

We examined process evaluations of interventions included in our review of outcome evaluations. Six reports of four separate studies were included. These employed various research methods, most frequently drawing on quantitative data collected from students and/or teachers. These reported positively on intervention feasibility, fidelity, reach and acceptability. The single study that examined context suggested that it was important, facilitating implementation when this built on schools' existing ethos and when senior staff championed the intervention.

Research question 4: multilevel studies

Multilevel studies measure outcomes at the individual level and explain these in terms of school- and individual-level student characteristics. Unlike ecological studies they can disentangle the effects of school-level factors that can also be represented at the individual level. We included 42 reports of multilevel studies (drawing on a total of 34 different data sets) examining the health effects of school-level factors measured 'objectively' (i.e. not merely aggregating data from individuals from whom outcome data were

collected). We confined our narrative synthesis to 10 reports that adjusted for key potential confounders and which did not overadjust for factors that might mediate school effects on health.

We found consistent evidence from studies of middle schools in the USA ($n=1$) and secondary schools in the UK ($n=3$) that schools with higher academic attainment and attendance than would be expected judging from the social profile of their students (i.e. a 'value-added' measure) had lower rates of substance use. The US study also reported that these schools have lower rates of group fighting and suggests that these school effects are generalisable to low-income, ethnic minority young people.

Findings on the influence of school policies were mixed. A German cross-sectional study of secondary schools reported that a complete smoking ban for students at or around school was associated with reduced smoking. However, a cross-sectional survey of secondary schools in the USA and Australia found no association between various forms of school smoking policies (including policies with constructive sanctions for students caught smoking) and any measures of student smoking. These differences between studies may reflect a 'ceiling' effect for the impact of smoking bans, which have already been widely implemented in US and Australian but not German schools. A cross-sectional study of Dutch secondary schools reported no associations between school policies on alcohol use at school or school sanctions and heavy drinking among students aged 12–16 years.

A cross-sectional study found that students in US middle schools with larger total campus and playground areas per student had higher rates of physical activity at school. A cross-sectional study of US high school students found that the number of unobservable/unsupervised places at school was associated with some measures of the use of alcohol and marijuana in school in the previous 12 months but not overall use in the past year. Finally, a cross-sectional study reported that the following school-level factors were not associated with alcohol use among students aged 13–14 years in rural schools in the USA: whether eighth graders are located within the same school as high school students or are in separate schools, school size and pupil–teacher ratio.

These multilevel studies provide little evidence on the impact of schools on health inequalities. Only one well-adjusted study of school effects examined subgroup effects (defined in terms of baseline health behaviour rather than socioeconomic status) and found no significant differences.

Research question 5: qualitative studies

In total, 21 qualitative studies were synthesised to explore the processes through which school-level influences might occur. Various pathways were identified. First, aggressive behaviour and substance use may be students' active responses to schools when they feel educationally marginalised or unsafe, which may in turn exacerbate disengagement and anxiety. Second, positive teacher–student relationships appear to be critical in promoting student well-being and limiting risk behaviour, although certain aspects of schools' organisation may have the potential to undermine these. Third, because of having so little involvement in decision-making in schools, students can fail to develop what social control theory defines as a 'stake' in their school, thus increasing the likelihood that they will instead look for a sense of identity and social support through health-risk behaviours. Fourth, students' lack of satisfaction with school can cause them to seek sources of 'escape', either through heavy drug use and drinking, or by leaving school at lunchtime or for longer unauthorised spells.

Conclusions

We focused on how schools are managed, designed and built and provide learning and teaching, pastoral care and discipline. There is evidence for the potential of school environment interventions addressing these to promote health, but the evidence is far from definitive. Five outcome evaluations examined interventions encouraging staff/students to build a stronger sense of community and/or better interpersonal relations in a range of US/UK school settings. These evaluations generally reported benefits for measures

related to emotional health and aggression. Two evaluations assessed interventions modifying American middle schools' food/physical activity environments and empowering students' involvement in this, reporting benefits for physical activity measures but not for diet. Process evaluations positively reported on interventions' feasibility, fidelity, reach and acceptability. To develop a fuller picture of the effects of school environment interventions, the results of our own review should be read in conjunction with those of the Cochrane review of HPS interventions, which include school environment alongside curriculum and parent/community components.

Outcome and process evaluations were subject to methodological limitations, and were not informed by nor aimed at testing any of our review theories. Most of the interventions employed multiple components addressing different aspects of schools' organisation and practice and so they do not lend themselves to testing specific hypotheses. However, the evidence from these lends broad support to each of our three primary hypotheses arising from the social development model (regarding the importance for health of participation in school activities), social capital theory (regarding the effects of trusting relationships) and the theory of human functioning and school organisation (regarding the importance of eroding rigid social boundaries between staff and students and how more student-centred framing of activities will enable better health outcomes).

The multilevel studies provide greater insights regarding our review hypotheses, most notably regarding the theory of human functioning and school organisation, which several studies explicitly aimed to test and provided evidence for.

The meta-ethnography of qualitative studies also supported the theory of human functioning and school organisation, suggesting that a lack of safety at schools, weak student-staff relationships, lack of student participation in decisions and educational disengagement may harm student health.

We have concluded that, although existing interventions suggest the potential for school environment interventions to promote young people's health, the evidence base is currently far from definitive. There is a need for better-conducted RCTs, studies outside the USA and studies on interventions focused on outcomes other than violence, healthy eating and physical activity. The multilevel studies and qualitative evidence reviewed have suggested potential new foci for intervention studies, such as interventions addressing student engagement, attainment and attendance, student participation in decisions, and the school physical environment. More trials are also needed to improve the evidence base concerning interventions addressing school community building and interpersonal relationships, particularly in secondary schools and outside the USA. RCTs of playground improvements are also required.

Funding

The National Institute for Health Research Public Health Research programme.

Chapter 1 Background

The importance of the school environment

Young people in the UK are among those with the worst health in Europe, and there are marked health inequalities, with considerable implications for later health and economic costs.^{1,2} There are increasing suggestions that seemingly separate outcomes, such as substance use, violence and sexual risk, are interlinked, requiring common intervention strategies.^{3,4} Health education delivered through the school curriculum and aiming to improve knowledge, develop skills and modify peer norms is now well established in schools, addressing smoking, drinking, drug use, sexual behaviour, physical activity and diet. However, systematic reviews suggest that such interventions have disappointing results.⁵⁻¹¹

A complementary approach to curriculum-based health education is to change the school environment to promote health and well-being. Although, traditionally, research has focused on individual and family risk and protective factors, there is increasing recognition that young people's health can be influenced by broader social factors.^{12,13} A large body of educational research has explored 'school effects' on attainment and other outcomes. Such research originated in the 1970s with the work of Michael Rutter *et al.*,¹⁴ which questioned the previous assumption that young people's educational attainment was determined mainly by their social background, with schools having little or no effect.¹⁵ Rutter and others' work suggested that schools differ in student attainment and that factors such as strong leadership, student involvement, high expectations and frequent evaluation and praise appeared to explain some of these differences.¹⁶⁻²⁰ Different schools were described as having a different 'ethos', referring to the sets of values, attitudes and behaviours distinguishing one school from another.^{20,21} As well as their social environments, schools differ in their physical environment, such as cleanliness, lighting, ventilation and aesthetic appeal, which may have important consequences for students' engagement and learning.²²

The school environment may also have profound effects on students' emotional and mental health, and opportunities to choose healthy lifestyles, a point first suggested by early studies showing significant differences in rates of health-related behaviours and outcomes between schools.²³ Rather than treating schools merely as sites for health education, 'school environment' interventions aim to modify how the school social and physical environment influences health. School environment interventions can address health directly, for example by modifying school policies (e.g. on smoking),²⁴ improving catering²⁵ or encouraging staff and students to walk or cycle to school.²⁶ Other actions aim to address factors such as disengagement and a lack of social support, which are risk factors for multiple adverse outcomes.^{27,28} Such actions include increasing student participation in decision-making and providing staff with training on how to re-engage disaffected students. These interventions take a 'socioecological'²⁹ approach to promoting health, whereby health is understood to be influenced not only by individual characteristics and behaviours but also by the wider social, cultural and economic context.

An important influence on the development of school environment interventions has been the World Health Organization's (WHO) framework for Health Promoting Schools (HPS).³⁰ This was influenced by WHO's Ottawa Charter, which recognised the limited effectiveness of health education alone in promoting health.³¹ The framework for HPS called for health to be promoted through the whole school environment, and not just through 'health education' in the curriculum. A 'health promoting school' aims to promote lifestyles conducive to good health; provide an environment that supports and encourages these lifestyles; and enable students and staff to take action for a healthier community and healthier living conditions.³² In England, the previous government developed a National Healthy Schools Programme, informed by the WHO framework, to which all schools were to sign up; however, the current government has ceased providing national funding for this programme and rendered it optional for schools to participate.

The HPS model has been described in various ways in different documents.^{33–40} Although some interventions are explicitly labelled as adopting a HPS model, others do not use this name but nonetheless are implicitly based on the same principles. In the USA, this approach is commonly referred to as a Coordinated School Health Program (CSHP). Both HPS and CSHP require change in three areas of school life: (1) formal health education curriculum, (2) environment and ethos and (3) links with the wider community.

Some approaches to HPS have been rigorously evaluated, but many have not.³⁴ Other trials have evaluated interventions that aim to modify the school environment to promote health but which are not explicitly informed by the HPS framework.

Rationale for the systematic review

Evidence concerning the effects of school environment interventions has not been comprehensively synthesised and reviews that have examined these interventions are now quite old. Existing reviews have also focused on interventions combining school environment change with curriculum components and so cannot assess the contribution of environmental change to any health effects detected. A decade-old systematic review of HPS interventions identified only 12 studies, four of which were randomised trials. It concluded that HPS interventions are promising, especially for promoting healthy eating, reducing bullying and improving mental and social well-being.³⁴ Other reviews of school-based interventions have similarly examined interventions addressing the school environment alongside other forms of intervention such as classroom curricula and counselling. A meta-analysis of school-based interventions to address a range of problem behaviours concluded that such interventions were effective in reducing alcohol and drug use,⁴¹ a point echoed in a more recent systematic review focused solely on whole-school interventions to prevent drug use.⁴² Because existing reviews have examined interventions combining school environment with other components they cannot assess the specific effects on health of environment changes. Furthermore, no evidence syntheses have been carried out on the effects of school environment interventions in important areas such as sexual health, alcohol or smoking.

There has also been no synthesis of evidence on the school environment intervention process. Process evaluations examine the planning, delivery and receipt of school environment interventions and how these are influenced by local context, and are useful for informing decisions about the wider implementation of interventions.^{43,44} Process evaluations can help explain how and under what conditions an intervention works and so can form a useful complement to randomised controlled trial (RCT) examination of whether or not and for whom an intervention works.⁴⁵

A further gap concerns the synthesis of evidence on the health effects of the school social and physical environment in the absence of specific interventions. Examining the impacts of such school-level factors on health outcomes is now a growing field of public health research that merits synthesis.⁴⁶ Although such observational studies provide less certain causal inference than experimental studies, those aiming to minimise confounding and other sources of bias could be used to identify promising areas for future intervention studies. This is important because, to date, school environment intervention studies appear to have addressed only some aspects of the school environment and neglected others, such as school leadership and approaches to learning.

An early review of the effects of anti-smoking policies on student smoking concluded that there was some evidence that these were effective; however, the review was hampered by its non-systematic design and admission of ecological alongside multilevel studies. Multilevel studies, unlike ecological studies, enable proper examination of how features of the school as an institution as opposed to the compositional features of the student body affect student health outcomes.⁴⁷ The review by Aveyard *et al.*⁴⁸ acknowledged the importance of multilevel evidence; however, it concluded that, although smoking prevalence differed markedly between schools, it was not yet possible to determine whether this was due to differences in student composition or to schools as institutions because of the poor methodology of studies. A particular

problem was that studies did not adequately adjust for the potentially confounding effects of the families and neighbourhoods from which students were drawn, and overadjusted for factors that might actually mediate school-level effects on smoking, such as student attitudes to school and peer behaviours. This review used a simple set of database search terms, which may have made for an insensitive strategy, and the review authors noted that many of the articles included were found through the reference lists of included reports rather than from the database searches. Another review of multilevel studies of school effects on a range of student outcomes, including health as well as academic performance, involved a yet more rudimentary search strategy and no prespecified methods of quality appraising and synthesising studies.⁴⁹

Finally, qualitative research has also been used to explore how staff and students perceive their school environment, and the processes they see as influencing health.⁵⁰ This evidence would also be useful in informing future school environment interventions but remains unsynthesised.

Our review aimed to address these gaps. It was conducted in close collaboration with colleagues undertaking a Cochrane review that is updating the decade-old review of HPS interventions.³⁴ This Cochrane review focuses on interventions addressing each of the following areas: school curriculum; environment or ethos of the school; and links with parents/the wider community.⁵¹ Our review instead focused only on school environment interventions that lack a health education curriculum component. This was a pragmatic means to provide our review with a focus distinct from that of the Cochrane review, but also allowed us to examine whether or not it is possible to attribute health effects to changes to schools' social and physical environments.

Chapter 2 Aim and research objectives

The initial, overarching purpose of this systematic review was to synthesise evidence relating to the effects of interventions addressing, and school-level measures of, schools' social and physical environments on the health and well-being of students and staff.

The research objectives and hypotheses were refined across two stages. In the first stage we developed broad research questions (RQs) geared towards developing a map of evidence and theories related to the review. These encompassed all aspect of schools' social and physical environments and the health and well-being of both students and teachers. These data were then presented to stakeholders (academics, people working in policy and practice and young people) whom we consulted with to help focus the review. We refined our research objectives in light of these consultations and in stage 2 focused specifically on student health and defined the school environment more narrowly in terms of how schools are organised/managed, how they teach, provide pastoral care and discipline students, and/or the school physical environment. We conducted five in-depth reviews of the evidence corresponding to the following RQs.

Research question 1

What theories and conceptual frameworks are most commonly used to inform school environment interventions or explain school-level influences on health? What testable hypotheses do these suggest?

Research question 2

What are the effects of school environment interventions (interventions aiming to promote health by modifying how schools are organised and managed; or how they teach, provide pastoral care to and discipline students; and/or the school physical environment) that do not include health education or health services as intervention components and which are evaluated using prospective experimental and quasi-experimental designs, compared with standard school practices, on student health [physical and emotional/mental health and well-being; intermediate health measures such as health behaviours, body mass index (BMI) and teenage pregnancy; and health promotion outcomes such as health-related knowledge and attitudes] and health inequalities among school staff and students aged 4–18 years? What are their direct and indirect costs?

Research question 3

How feasible and acceptable are the school environment interventions examined in studies addressing RQ2? How does context affect this, examined through process evaluations linked to outcome evaluations reported under RQ2 above?

Research question 4

What are the effects on health and health inequalities among school students aged 4–18 years of school-level measures of school organisation and management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers, and/or the physical environment (measured using 'objective' data other than aggregate self-reports of the same individuals who provide data on outcomes), examined using multilevel quantitative designs?

Research question 5

Through what processes might these school-level influences occur, examined using qualitative research?

Protocol

The review protocol is available in *Appendix 7*. The published version can be freely accessed from the BioMed Central website (www.biomedcentral.com/1471-2458/11/453): Bonell C, Harden A, Wells H, Jamal F, Fletcher A, Petticrew M, *et al*. Protocol for systematic review of the effects of schools and school-environment interventions on health: evidence mapping and syntheses. *BMC Public Health* 2011;**11**:453.

Chapter 3 Report organisation and terminology

The report is organised according to the two stages of the research project. Stage 1: identifying and describing the references (see *Chapter 5*) presents the RQs, methods and findings of the evidence and theory map, and the stakeholder consultations. Stage 2: in-depth synthesis (see *Chapters 6–10*) presents the methods, results and discussions for the in-depth reviews for each RQ. *Chapter 6* presents the in-depth review of the theories; *Chapter 7* presents the in-depth review of the outcome evaluation studies; *Chapter 8* presents the in-depth review of the process evaluation studies; *Chapter 9* presents the in-depth review of the multilevel studies; and *Chapter 10* presents the in-depth review of the qualitative studies. Each of these chapters lists the RQ investigated, explains our methods, gives an overview of the included reports and presents the results and discusses these in relation to the RQ at hand. In *Chapter 11* we develop an overall synthesis in which we assess the primary and secondary review hypotheses developed in *Chapter 6* in relation to the empirical evidence presented in *Chapters 7–10*. *Chapter 11* also considers the strengths and weaknesses of the review, provides a summary of our findings and suggests implications of our review.

We use the term 'report' to refer to written publications included in the review. We use 'study' or 'data set' to refer to the research from which these arose. We use 'reference' to mean records of study reports included in the evidence map. 'Statistically significant' is used to indicate $p < 0.05$, except where otherwise indicated.

Chapter 4 Data management

We used EPPI-Reviewer 4 (ER4; Evidence for Policy and Practice Information and Co-ordinating Centre, Social Science Research Unit, Institute of Education, University of London, London, UK) to support the management and analyses of the references found and the data extracted for all stages of the review.⁵² ER4 is a web-based systematic review program that supports the review process: downloading of bibliographic citations, application of inclusion and exclusion criteria, recording and storing free text and categorical and numerical data, and conducting statistical and qualitative synthesis. This specialist program also incorporates functions for comparing the independent assessments of reports from two or more reviewers. Therefore, ER4 helped to assure quality in our review and facilitated transparency and auditability.

Stage 1 Identifying and describing the references

Chapter 5 Evidence map, theory map and stakeholder consultations

Aim and research questions

The purpose of the map of evidence and theory and stakeholder consultations was to identify references that are potentially relevant to our review questions; to assess the nature of the references; and to refine our review questions for stage 2. The RQs for this initial mapping stage focused on all aspects of schools' social and physical environment and therefore were broader than the refined questions that we finally examined in our in-depth reviews.

Research question 1

What theories and conceptual frameworks are most commonly used to inform school environment interventions or explain school-level influences on health? What testable hypotheses do these suggest?

Research question 2

What are the effects of school environment interventions (interventions aiming to promote health by modifying the school's physical, social or cultural environment through actions focused on school policies and practices relating to education, pastoral care, sport, extracurricular activities, catering, travel to and from school and other aspects of school life) evaluated using experimental and quasi-experimental designs, compared with standard school practices, on health (physical and emotional/mental health and well-being; intermediate health measures such as health behaviours, BMI and teenage pregnancy; and health promotion outcomes such as health-related knowledge and attitudes) and health inequalities among school staff and students aged 4–18 years? What are their direct and indirect costs?

Research question 3

How feasible and acceptable are school environment interventions? How does context affect this?

Research question 4

What are the effects of other school-level factors on health and health inequalities among school staff and students aged 4–18 years, examined using multilevel and ecological (school) designs?

Research question 5

Through what processes might these school-level influences occur?

Methods

Database searching

Electronic databases searched

A total of 16 bibliographic databases were searched between 30 July 2010 and 23 September 2010, with no limits on language or date:

- Australian Educational Index
- British Educational Index
- CAB Health (part of CAB Abstracts) – now known as Global Health
- The Campbell (C2) Library
- Cumulative Index to Nursing and Allied Health Literature (CINAHL)

- Cochrane Central Register of Controlled Trials
- EMBASE
- Education Resources Information Center (ERIC)
- Health Management Information Consortium (HMIC)
- International Bibliography of the Social Sciences (IBSS)
- MEDLINE
- PsycINFO
- Social Policy and Practice (includes ChildData and Social Care Online)
- Social Sciences Citation Index (Web of Knowledge)
- Sociological Abstracts
- Dissertation Abstracts/Index to Theses.

EconLit and Public Affairs Information Services (PAIS) were also investigated, but trial searches produced no new material.

Search terms

A broad approach to database searching was used in stage 1 given the cross-disciplinary nature of the review, the wide range of study designs to be included and the variability with which references were indexed in bibliographic databases. A sensitive search was undertaken using a large number of natural-language phrases. The search terms were used to develop *core searches* that included the most relevant terms and in which references were to be scanned carefully, examining the full title/abstract in detail for inclusion; and *non-core searches* in which a broader set of 'non-core' (or marginal) terms were applied and scanning for inclusion was to be carried out slightly more rapidly (although in practice both were scrutinised carefully). Some additional intervention terms were added to the key terms as a third searching phase.

Core search

- Setting (1) – school terms.
- Population (2) – child terms.
- Intervention/effect (3A) – key intervention/school-level effect terms.
- Outcomes (4) – broad range of health outcomes.
- Key phrases (5) – related to health and schools.

Search 1: Set 1 and Set 2 and Set 3A and Set 4 (setting/population and key interventions/effects and outcomes).

Search 2: Set 5 (HPS phrases).

Non-core search

- Setting (1) – school terms.
- Population (2) – child terms.
- Intervention/effect (3B) – other non-key terms related to intervention/school-level effect (general free text).
- Outcomes (4) – broad range of health outcomes.
- Key phrases (5) – related to health and schools.
- Key phrases (6) – simple phrases combined with Set 4 outcome terms.

Search 3: Set 6 and Set 4 (whole school phrases and outcomes).

Search 4: Set 1 and Set 2 and Set 3B and Set 4 (setting/population and key interventions/effects and outcomes).

Additional terms were added to Set 3B in the third phase of the search.

Inclusion and exclusion criteria

We included quantitative, qualitative and theoretical literature in the evidence map that theorised or empirically examined the effect of the school social and/or physical environment, interventions to address this and/or processes underlying these effects or interventions (not including the provision of health education or health-related goods or services) on the health or well-being outcomes of students (age 4–18 years) or staff. All references from the searches were uploaded into ER4 and duplicate references were removed (those scoring ≥ 0.85 on ER4's similarity score). Exclusion criteria were developed (*Table 1*) to remove irrelevant references and thereby identify relevant references. References were screened on title and abstract. A round of pilot screening was conducted by two reviewers on a sample of 200 abstracts to test and refine the criteria. The remaining references were divided between six reviewers (CB, HW, AH, CV, MP and FJ) and screened independently. After each reviewer had screened 2000 references, a random sample of 10% ($n=200$) were double screened by another reviewer to ensure consistency in applying the criteria. A threshold of <20 disagreements per 2000 references on whether to include/exclude was established.

It should be noted that, although these criteria were applied to most of the references, because of time constraints and the large number of references, those that were obviously to be excluded were marked 'exclude only' and not assigned an exclusion code.

Evidence map: coding references

Included references were descriptively coded based on title and abstract. Descriptive coding involved identifying the following characteristics of each study:

- relevance to RQs 1, 2, 3, 4 and/or 5
- type of research
- country where research was undertaken
- research design
- target population
- health topic examined
- level of school (e.g. high school, elementary school).

A round of pilot coding was conducted by four reviewers (CB, AH, HW and FJ) on a random sample of 40 references to ensure that the list of characteristics captured was comprehensive and relevant. Two reviewers (FJ and HW) double-coded the remaining references to ensure consistency in coding. Discrepancies were resolved by discussion.

As a result of the large number of references included at this stage, references were coded for the evidence map on the basis of title and abstract only. When screening references for inclusion in the evidence map we erred on the side of inclusion and so there were inevitably errors of overinclusion.

Theory map

A map of theories and conceptual frameworks used to inform school environment interventions or explain school-level influences on health was developed alongside the evidence map. We looked for theories while coding the first half of the references for the evidence map to obtain a broadly representative sample of theories. Those theories that were 'named' (e.g. 'social learning theory') and which were referenced in multiple references were identified. Summaries of the included theories were obtained through a Google search or were extracted from the original texts where they were first published.

Consultation with stakeholders

To refine the RQs and focus the review, we consulted with key stakeholders regarding the review topic and evidence and theory map.

TABLE 1 Exclusion criteria for stage 1: evidence and theory map

Exclusion criterion	Guidance
Exclude 1: general topic	The study is not about health/well-being or disease (including references solely focused on outcomes concerned only with education)
Exclude 2: setting and population	The study is not about the students or staff of schools (i.e. serving those aged 4–18 years)
Exclude 3: type of report	The study does not report primary research, a review of research or a theory
Exclude 4: study focus	<p><i>Intervention (primary) references:</i> The intervention is neither mainly delivered on the school site nor concerned with travel to and from schools (extracurricular interventions were included unless excluded based on any of the criteria below); neither about an intervention aiming to promote health/well-being or prevent disease nor reporting on the health/well-being outcomes of an intervention; involves only health education, information or counselling (regardless of who delivers this), school nursing, clinics or health checks, or health-related goods (medication, contraception, micronutrients, etc.), but interventions concerning school catering, sport or active transport would be included; and targeted only to some students on the basis of health-related needs (but interventions targeted on the basis of educational or social but not health needs would be included)</p> <p><i>Non-intervention (primary) references:</i> The study is not related to the effects of the school environment/school-level factors on health/well-being. We excluded reports comparing health outcomes between individuals with different educational experiences or attitudes because such references cannot be used to infer school-level effects</p> <p><i>Reviews and theoretical references:</i> The study is not a review or theoretical paper with a focus on the school environment, interventions addressing this or school-level effects</p>
Exclude 5: study type ^a	<p><i>Intervention (primary) references:</i> The study is not an empirical outcome evaluation or process evaluation reporting on school environment intervention effects on health and/or cost, economic and econometric references examining school environment interventions</p> <p><i>Non-intervention (primary) references:</i> The study is not empirically examining school environment influences on health/well-being. If the study is a quantitative study it will be excluded if it is not reporting on school-level variables (but multilevel analyses including school-level analyses would be included); it is reporting only on school-level measures of students' social (e.g. socioeconomic status) or demographic (e.g. ethnicity) characteristics or students' social networks (but references examining student–staff relationships would be included); or it is reporting only on school-level measures of health education (regardless of who delivers this), school-based clinical health services or interventions targeted on the basis of health-related needs. If the study is a qualitative study it will be excluded if it is not reporting on the process by which schools might influence health</p> <p><i>Theoretical references:</i> The study does not propose an abstracted, generalisable way in which features of schools are causally related to student/staff health. In other words, include only literature describing/explaining the theories and conceptual frameworks that are used to inform school environment interventions or explain school-level influences on health</p> <p><i>Reviews:</i> The study is not a systematic review</p>

^a Studies that we identified evaluating interventions that involved curriculum components alongside school environment components were passed on to our Bristol/Cardiff colleagues. Our Bristol/Cardiff colleagues similarly passed on studies involving school environment components, but not curriculum components, to us.

Policy, practice and research

We presented the findings of the evidence map and theory synthesis to people working in policy ($n=3$), practice ($n=1$) and research ($n=2$) on 1 April 2011. These individuals were purposively selected to ensure expertise regarding young people's health and education and generate diversity according to sector. Based on the evidence map we engaged in semistructured in-depth discussions about:

- defining 'school environment interventions'
- determining the usefulness of theories in informing school interventions and explaining school-level effects

- establishing priorities for the review in terms of types of interventions, health outcomes and theories of interest.

The evidence map was presented by CB and the session was chaired and facilitated by AH. Discussion lasted just over 2 hours with notes being taken by HW.

Young people

We consulted with an existing group of young people, brought together to advise on the conduct of public health research. We met with DECIPHer (Development and Evaluation of Complex Interventions for Public Health Improvement)'s Public Involvement Advisory Group, called ALPHA (Advice Leading to Public Health Advancement), on 25 September 2010 and again on 27 May 2011. We originally intended to consult with young people once on the evidence map and once on the draft final report, but we decided to front-load young people's participation at the start to ensure that we examined areas of priority to them. This group includes young people aged 14–19 years from across south Wales.

The first consultation was conducted at project inception when reviewers were first developing the protocol. The purpose of this consultation was to find out what the terms 'health' and 'well-being' meant to young people and to elicit their perspectives on how schools might impact on their health and well-being. A total of 13 young people participated in a face-to-face semistructured consultation that lasted just over 1 hour and it was facilitated by two researchers (AH and RL), with oversight from one youth worker. Notes were taken by AH.

For the second consultation we presented findings of the evidence map and theory synthesis to the group. The purpose was to engage in a prioritisation exercise to find out which of the health outcomes identified in the evidence map young people found most relevant to their experiences. A total of 13 young people from the group participated in the face-to-face semistructured consultation facilitated by two researchers (AH and FJ), with oversight from one youth worker. Notes were taken by FJ.

Online consultations through a social networking site (<http://groups/youngpeopleinresearch>; note that this website is no longer active) supplemented the face-to-face consultations. ALPHA members were invited to join the online group and provide any further views on the questions elicited from the face-to-face consultations, but this resulted in minimal additional data.

Results

Flow of literature: from database searching to evidence map

A total of 82,775 references were retrieved from the database searching. Of these, 20,446 were identified as duplicates: either 'exact' matches ($n=19,132$) or very close matches ($n=1314$). The remaining 62,329 references were screened on title and abstract and 61,185 (98.2%) were excluded. In total, 1144 references were included in the evidence map (*Figure 1*).

Literature from additional searches

Eight reports were identified for the stage 2 in-depth syntheses from additional searches (citation chasing and contacting authors and collaborators). One additional multilevel study of school effects on health was identified by contacting authors of included reports.⁵³ Two intervention outcome evaluation reports were referred to us by our Cochrane review collaborators.^{54,55} Two additional theory references and one intervention outcome evaluation were identified by reference sifting.^{56–58} Two additional intervention outcome/process evaluation reports were included as suggested by CB.^{59,60} The reasons why these references were not captured in the database search were because they were published after our database search date,^{53,59} because of the reference type⁵⁴ (conference paper) and/or because they were lacking relevant key

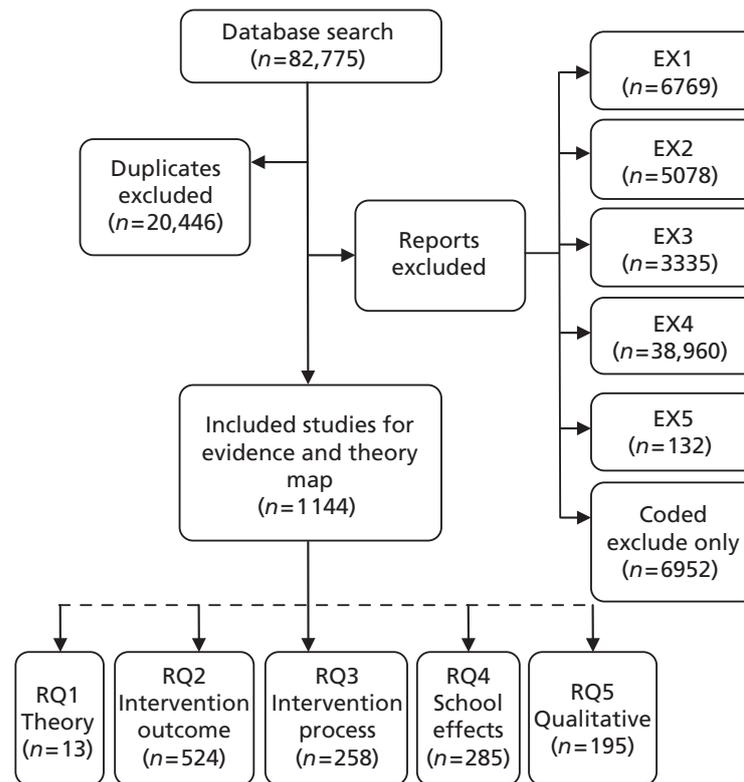


FIGURE 1 Flow of literature stage 1: evidence and theory map.

wording.^{55,58} Reports identified from additional searches are not presented in the results of the evidence map or flow of literature diagram because the additional searches were conducted after the map was produced.

Evidence map

The 1144 references were descriptively coded based on title and abstract to identify relevant characteristics of references. Because the references were coded for inclusion on title and abstract only, there are inevitably errors of overinclusion. Nonetheless, the evidence map provides a useful overview of the available evidence to inform our decisions about what references to prioritise for in-depth review in stage 2.

Relevant research question(s) of references included in the evidence map

Figure 2 indicates to which RQs references might be relevant. The total number of references, displayed in this figure, does not equal the total number of included references as categories were not mutually exclusive.

Types of research

The vast majority of the references were coded as primary research ($n=1088$). Very few systematic and other literature reviews ($n=68$) were identified and even fewer stand-alone theory/conceptual references ($n=9$).

Country where research was undertaken

Figure 3 provides the distribution of research for the top 10 countries where research was conducted. Some references did not report the country of research in the title or abstract ($n=433$). The total number of references, displayed in this figure, does not equal the total number of included references as categories were not mutually exclusive.

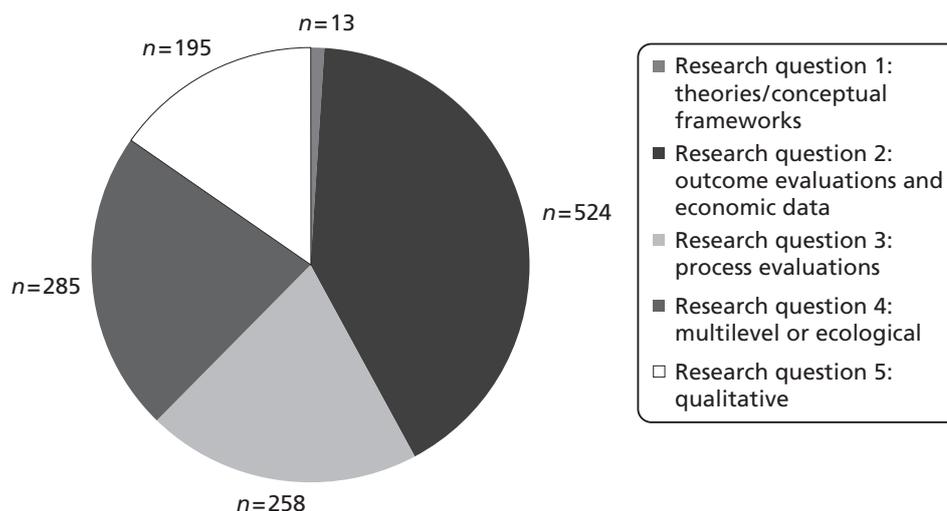


FIGURE 2 Potential RQ(s) addressed by references in the evidence and theory map. RQ1: literature describing/explaining the theories and conceptual frameworks that were used to inform school environment interventions or explain school-level influences on health. RQ2: evaluation references reporting on school environment intervention effects on health, as well as cost, economic and econometric references examining the costs of school environment interventions. RQ3: process evaluations of school environment interventions. RQ4: multilevel or ecological (school) references examining school-level influences on health. RQ5: qualitative references exploring the process by which school-level factors might influence health.

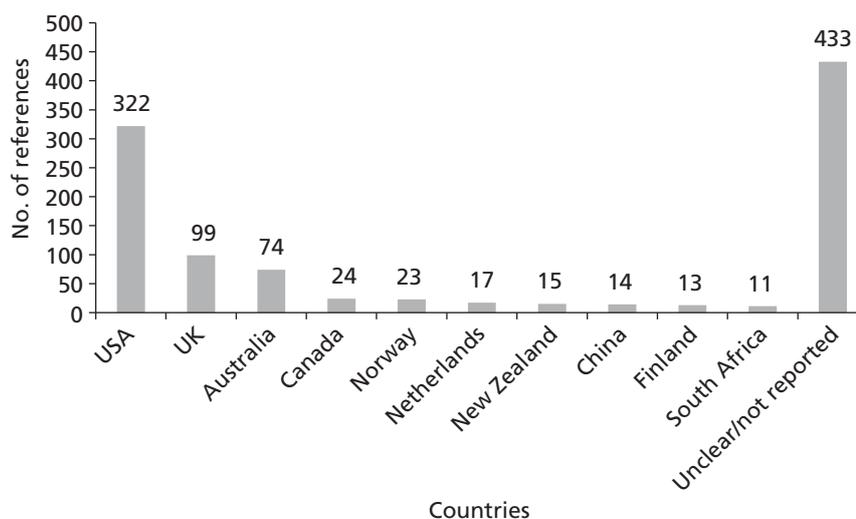


FIGURE 3 Countries of primary research of studies included in the evidence map.

Research design

For references identified as outcome evaluation references reporting on the effects of school environment interventions on health, the primary research design was coded as follows:

- RCTs ($n=143$)
- non-randomised comparison groups ($n=111$)
- before/after with no comparison groups ($n=74$)

- single cross-sectional survey ($n=20$)
- other [$n=7$: potential quasi-experimental ($n=1$), participatory action research project ($n=1$), quantitative data evaluation ($n=1$), focus groups ($n=2$), meta-analysis ($n=1$) and interrupted time series ($n=1$)].

The research design was not clear from the title/abstract for 99 of the outcome evaluation references. Only nine references were identified as potentially including cost data or having conducted an economic analysis; the vast majority did not report this information in the title/abstract.

For references identified as potential multilevel or ecological (school) references examining school-level influences on health, the design was coded as follows:

- single cross-sectional surveys ($n=198$)
- longitudinal cohort or repeat cross-sectional ($n=42$)
- other [$n=4$: prospective diary design ($n=1$), in-depth interview and focus group ($n=1$), policy analysis ($n=1$) and observational study ($n=1$)].

The remaining multilevel/ecological study references ($n=39$) did not clearly report the research design in the title/abstract.

Target population and health topics examined

Students were the target population studied for nearly all of the references included in the evidence map ($n=1093$), with only 52 titles/abstracts mentioning staff as the target population.

Most of the health topics identified were about violence, bullying and/or harassment ($n=278$), eating/drinking (non-alcoholic; $n=270$) or physical activity ($n=257$). *Figure 4* provides the distribution of the different health topics examined among included references. The total number of references displayed in the figure does not equal the total number of included references as categories were not mutually exclusive.

School level/grade level reported

Most titles/abstracts did not report the school/grade level studied ($n=519$). Of those that did, the majority of the research was conducted at elementary/primary schools ($n=167/125$) or high/secondary schools ($n=126/117$). *Figure 5* provides the distribution of the different school/grade levels examined. The total number of references, displayed in the figure, does not equal the total number of included references as categories were not mutually exclusive.

Aspect of the school examined

Figure 6 provides the distribution of the different aspects of schools examined among included references. Most reports focused on school management or policies ($n=273$), catering or vending ($n=216$) and sport or active transport ($n=182$). We were unable to determine the aspect of the school examined for about one-quarter of the references based on the title/abstract alone (not clear $n=204$; not reported $n=157$). The total number of references, displayed in the figure, does not equal the total number of included references as categories were not mutually exclusive.

Theory map

A total of 12 theories/models were identified from the theory map. These include anomie theory, attachment theory, differential association theory, ecological systems theory, educational transmission of class theory, the health belief model, social cognitive theory, social control theory, the social development model, the social learning model, strain theory and the theory of reasoned action.

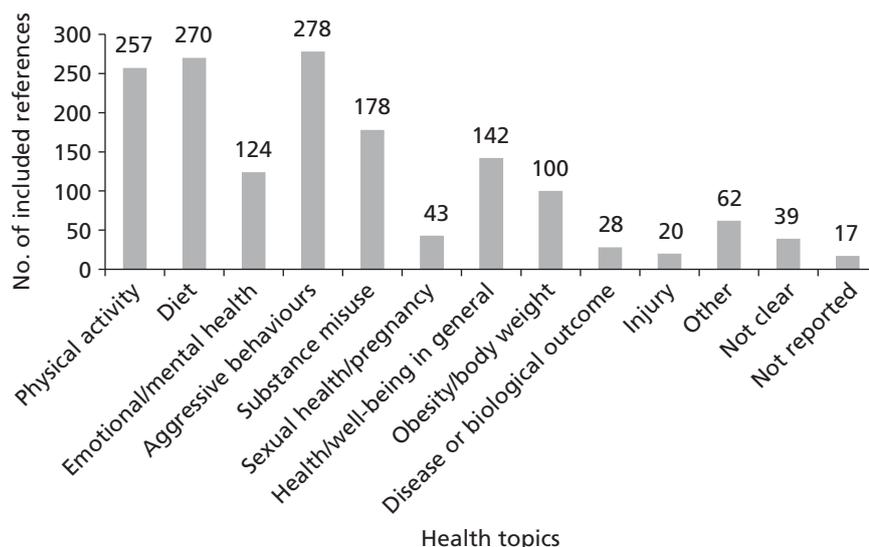


FIGURE 4 Health topics of the references included in the evidence map.

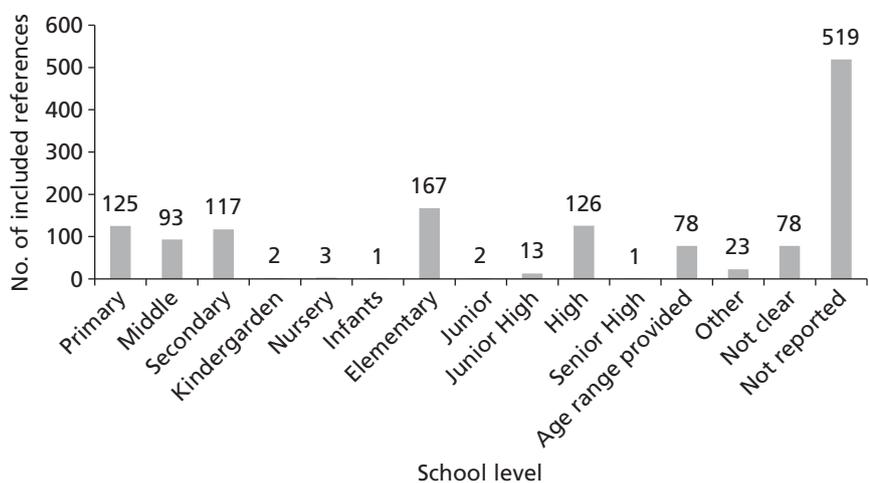


FIGURE 5 School/grade level of the references included in the evidence map.

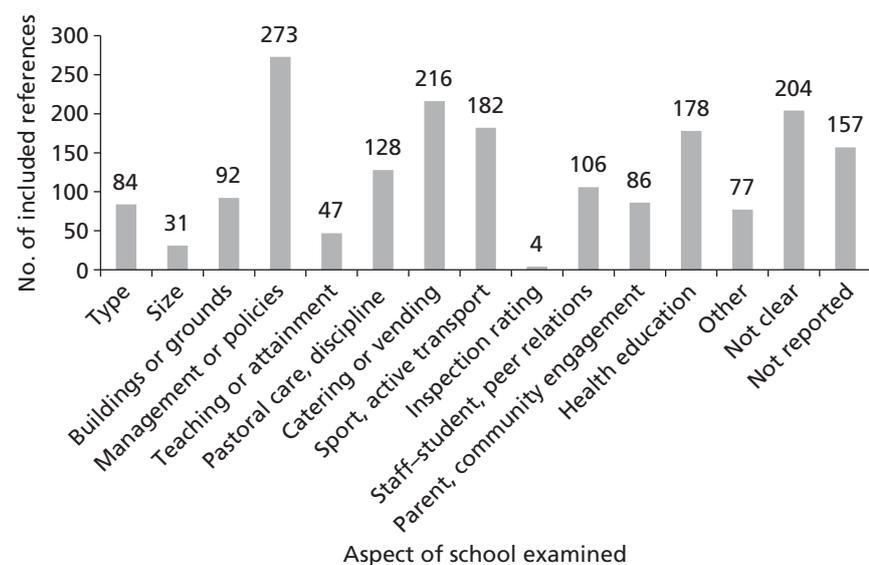


FIGURE 6 Aspect of the school examined in the references included in the evidence map.

The theories suggest the potential importance of school-level determinants concerning or interventions addressing:

- how schools structure norms (anomie theory)
- relationships between staff and students (attachment theory, social learning theory)
- the roles and opportunities that schools give to or withhold from students (social development model, strain theory)
- teaching and learning (educational transmission of class theory)
- school rules (social control theory)
- health education (health belief model)
- a combination of these (ecological systems theory).

Stakeholder consultation

The stakeholder consultation with policy-makers, teachers and academics suggested that we needed to define the school environment more clearly because otherwise it might be assumed to be the physical environment only. The consultation group suggested that, based on the presentation of the evidence map, we narrow our focus to:

- policy and management (policies, systems)
- social relationships (including staff–student, student–student and staff–staff relationships)
- student culture (sense of connection, engagement and aspiration)
- staff culture (values, vision, priorities, ethos, leadership)
- physical environment (school grounds).

These stakeholders suggested that they would most value a synthesis of evidence on the effects of schools' 'core business' on student health in terms of (1) learning and teaching, (2) pastoral care and (3) discipline policies/practices. The mental health of teachers was also considered important by the stakeholders with whom we consulted.

Young people told us that being healthy and well meant feeling safe and secure, having personal confidence, feeling self-assured and having the support of friends and family. Young people suggested that schools affect their health and well-being in various ways and emphasised the importance of:

- class size (e.g. large classes may mean less personal support, although some young people also thought small classes could be stifling)
- staff attitudes (e.g. having to spend time with teachers in a 'bad mood' was unhealthy)
- choice and empowerment at school (e.g. having a say in the running of schools)
- social class composition (e.g. students from poorer backgrounds may feel or be made to feel out of place)
- socialising (e.g. making friends, meeting people from different backgrounds)
- teaching and learning (e.g. 'making you smarter', 'opening your mind')
- hygiene (e.g. 'disgusting toilets')
- school meal options and prices (e.g. healthy food is often more expensive).

Young people thought the following were the most important things that a school could do to improve student well-being:

- reduce class sizes
- foster a positive attitude in teachers and good relationships between staff and students
- focus less on 'league tables' and more on 'learning for learning's sake'
- increase opportunities for students to focus on what they are interested in or good at
- provide more sources of social support for students.

The young people we consulted with cited the following as important outcomes: good relationships (especially with teachers), anxiety, self-image and 'overachieving'. This informed our decision to focus on aspects of schools' 'core business' of teaching, pastoral care and discipline. Based on these consultations we decided not to focus on school activities such as extracurricular activities, catering and vending of food and drinks, physical education (PE) and active transport to and from school.

Implications for stage 2: in-depth review

The findings from the evidence and theory map as well as the stakeholder consultations with young people and people working in policy, practice and research suggested that the most important school environment determinants and interventions to focus on concerned those relating to how schools are organised and managed and how they deliver teaching, pastoral care and discipline, as well as the physical environment of schools. Therefore, we focused on aspects of schools' organisation and management, teaching, pastoral care and discipline and the physical environment that may influence student health outcomes. We chose not to focus on catering, PE, extracurricular activities or active transport to and from school. Our decision not to focus on school environment interventions involving changes to school catering, PE and extracurricular activities was informed by a view that these areas are already well synthesised.^{25,61} We also decided to review only references focused on student health and not teacher health despite this being recommended by some stakeholders, to make our review manageable.

Research questions for stage 2: in-depth review

Research question 1

What theories and conceptual frameworks are most commonly used to inform school environment interventions or explain school-level influences on health? What testable hypotheses do these suggest?

Research question 2

What are the effects of school environment interventions (interventions aiming to promote health by modifying how schools are organised and managed, or how they teach, provide pastoral care to and discipline students, and/or the school physical environment) that do not include health education or health services as intervention components and which are evaluated using prospective experimental and quasi-experimental designs, compared with standard school practices, on student health (physical and emotional/mental health and well-being; intermediate health measures such as health behaviours, BMI and teenage pregnancy; and health promotion outcomes such as health-related knowledge and attitudes) and health inequalities among school students aged 4–18 years? What are their direct and indirect costs?

Research question 3

How feasible and acceptable are the school environment interventions examined in references addressing RQ2? How does context affect this, examined using process evaluations linked to outcome evaluations reported under RQ2 above?

Research question 4

What are the effects on health and health inequalities among school students aged 4–18 years of school-level measures of school organisation and management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers, and/or the physical environment (measured using 'objective' data other than aggregate self-reports of the same individuals who provide data on outcomes), examined using multilevel quantitative designs?

Research question 5

Through what processes might these school-level influences occur, examined using qualitative research?

The chapters that follow describe how references in the evidence map were screened against a priori criteria to determine whether or not they were included in the in-depth reviews addressing each of the above questions.

Additional searches

Additional searches were conducted by screening the reference lists of all reports from the evidence map that were included in the in-depth review; contacting authors of included references for additional references; and asking Cochrane review collaborators for additional references. References published before June 2011 from the additional searches were considered for inclusion in stage 2.

Stage 2 In-depth synthesis

Chapter 6 Research question 1: theory synthesis

Research question

Which theories are cited in the literature and what hypotheses do they suggest for this review?

Methods

Inclusion and exclusion criteria

Thirteen stand-alone theory references were identified from the evidence and theory map as relevant to the theory synthesis. We included literature in the in-depth synthesis that theorised how the school social or physical environment (defined in terms of how schools are organised and managed, how they provide teaching, pastoral care or discipline, and schools' physical environment) affects students' health or well-being. An additional two references^{56,57} were sourced through reference checking. The full-text reports of these references were retrieved and the following exclusion criteria were applied independently by two reviewers (there were no discrepancies to be resolved):

- exclude reports that do not propose an abstracted, generalisable way in which core features of schools and school environment interventions are causally related as (1) a stand-alone theory, (2) a general theory of school health or (3) a theory addressing school influences on health
- exclude reports that are not written in English.

Quality assessment

The descriptions of the theories were extracted from included reports. We then obtained the original source of the theory and used this as a focus for quality assessing theories. The criteria for quality assessing theories are as follows:

- whether or not the constructs are well specified
- whether or not clear causal pathways are specified between constructs
- whether or not it was a simple theory/model
- whether or not it suggested which specific aspects of the school institution might influence health
- whether or not it is applicable to multiple health domains
- what the theory/model assumptions are
- whether these assumptions are implicit or explicit.

We developed these ourselves having searched for but not found existing criteria to assess the quality of theories. Our criteria were intended to determine which theories to use to inform the development of overall hypotheses for the review; these focus on the internal logic of each theory (well-specified constructs; clear pathways; simple; explicit assumptions) and its applicability to understanding school effects on health (which specific aspects of an institution influence health; applicable to multiple health domains). Some of these criteria were necessarily subjective, calling for researcher judgements, for example about whether or not a theory was simple (i.e. parsimonious). The quality assessment criteria were piloted on a random sample of two theories by two reviewers (CB and HW) before being applied by one reviewer (CB) and checked by another (HW), with any differences being settled by discussion.

The quality criteria were used to categorise theories as either primary or secondary theories, which in turn inform our primary and secondary review hypotheses. Theories were not excluded based on quality scores. We used these quality criteria to form a judgement about which theories to draw on to define our primary

and secondary review hypotheses. We did not simply require that a theory meet every criterion to be deemed 'primary' because our judgements were necessarily more subtle than this.

Data extraction

We extracted data related to the name of the theory, its originator, the year of origin, what constructs and pathways it involved, its disciplinary origins and whether it is linked to any higher-order or lower-order theories. The data extraction tool was piloted by two reviewers (CB and HW) on a random sample of two theories before being applied by one reviewer (CB) and checked by another (HW), with any differences being settled by discussion.

Synthesis

We summarised the primary and secondary theories in tables and used these to inform the development of hypotheses for our review, which we assess against the empirical evidence reviewed and report in *Chapter 11*.

Overview of included reports

Flow of literature

Of the 13 references identified from the evidence and theory map, only one⁶² was included. The other 12 references were excluded because they did not report a stand-alone theory ($n=9$), did not report a general theory of school health ($n=2$) or did not address school influences on health ($n=1$). Two reports identified through citation chasing were included.^{56,57} All empirical reports included in RQ2–5 were screened for reference to any theories. A total of 35 reports were identified through this search and were included in the in-depth synthesis. Thus, 38 reports, which reported on 24 theories, were included in the in-depth review (*Figure 7*).

Quality assessment

We judged that three of the 24 theories (ecological systems theory, ecological model of co-ordinated school health programmes and theory of structuration) did not have clear or narrowly specified constructs in the sense that these might be operationalised in empirical research, but that the rest ($n=21$) did. We judged that nearly all of the theories used a reasonably small number of components and a simple framework of inter-relations between them to understand potential school effects on health ($n=20$). Four were categorised as more complex (ecological systems theory, ecological model of co-ordinated school health programmes, human functioning and school organisation and the theory of triadic influence). We judged causal relations between constructs as being clearly specified in all but three of the theories (contagion theory, ecological systems theory and ecological model of co-ordinated school health programmes). Fourteen of the 24 theories were judged not applicable to multiple health domains: 11 applied only to delinquency outcomes, two to public behaviour only and one to violence only. Moral authority theory, social control theory, deterrence theory, the integrated perspective on delinquent behaviour and strain theory were all judged as relevant to understanding health outcomes that are associated with antisocial behaviour (ASB) only. These theories are not relevant to understanding other health outcomes such as healthy eating and physical activity.

We judged that three theories fully met the criterion of whether or not the theory suggests which specific aspects of the school might influence health. Another 10 partially met it.

Study characteristics

Twenty-four theories were identified in the in-depth theory synthesis. These were cited in a total of 38 reports^{53,56,57,62–96}. The theories most commonly cited in empirical reports were ecological systems theory^{68,75,77,78,80,83,85,90,96} ($n=10$), social control theory^{71,73,80,82,97} ($n=6$), social disorganisation theory^{70,79,88,89,97} ($n=5$), social learning theory^{71,76,81,94} ($n=4$), the theory of human functioning and school organisation^{53,64–66,95} ($n=5$) and social cognitive theory^{63,76,77,83} ($n=4$).

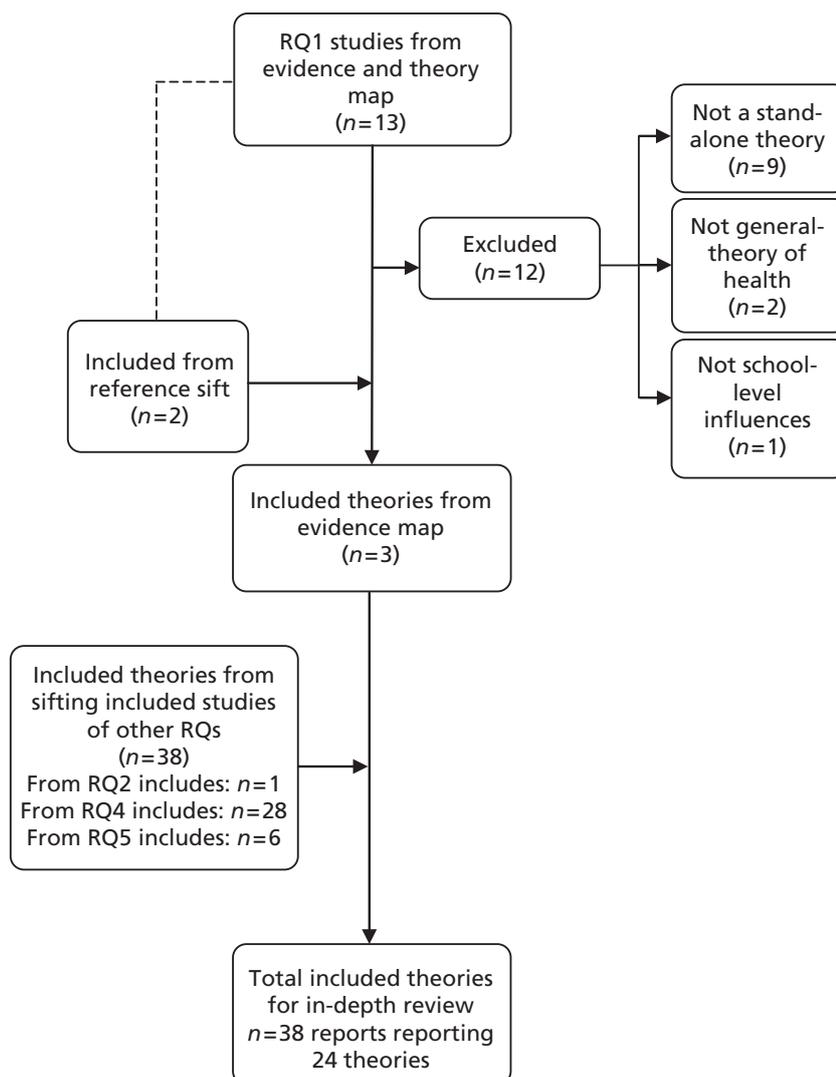


FIGURE 7 Flow of literature: theory synthesis (stage 2: in-depth synthesis).

Results

Primary theories that explain the mechanisms by which schools determine health

We identified three theories to be most useful, informing the primary hypotheses to use in interpreting findings from the review's empirical studies. These met all of our quality criteria, with the exception of the theory of human functioning and school organisation, which did not meet our criterion of simplicity. However, we judged that this theory was very strong regarding our other criteria and so included it.

Social development model

The social development model is a social psychological theory developed by Hawkins and Weis,⁹⁸ building on social learning and social control theories. This suggests that young people learn antisocial and prosocial patterns of behaviour from their immediate social environment by being provided with opportunities for involvement, opportunities to develop skills and reinforcements for actions. These processes build attachment to others engaged in these activities and, where these activities are prosocial, potentially build commitment to the conventional social order and conformity with social norms. This theory would lead us to expect that schools are more likely to foster attachment to the school and encourage healthy behaviours if they provide opportunities for students to participate in learning and institutional life; enable development of the skills necessary for such participation; and ultimately enable students to gain recognition and prosocial reinforcement for this.

Social capital theory

Social capital is conceived somewhat differently by different scholars; however, all conceptualise social capital as the product of social structures that facilitate group or individual actions.⁹⁹ The theory suggests that social networks characterised by reciprocity, trust and shared social norms will facilitate such actions, although whether such actions are health promoting or harming will depend on the specific nature of the norms. Theorists differ about whether social capital should be considered a property of groups, as Putnam¹⁰⁰ suggests, or individuals, as Bourdieu and Wacquant¹⁰¹ argued. They also differ over whether it should be considered primarily as an adjunct to economic capital among elites, as suggested by Bourdieu and Wacquant,¹⁰¹ or as a distinct resource available for use by all, as Coleman¹⁰² suggested.

Social capital was theorised by Coleman¹⁰² as being strongest when the networks involved are stable, enclosed and intergenerational and involve norms of reciprocal obligation. Informed by Granovetter,¹⁰³ Putnam¹⁰⁰ introduced the distinction between bonding and bridging social capital, the former being strong ties between similar individuals and the latter being weaker ties between more disparate individuals and groups. Granovetter¹⁰³ had earlier pointed to the importance of weak ties in communicating information and norms. Portes⁹⁹ has pointed to the potential for bonding social capital to lower aspirations and reducing individual autonomy.

Drawing from these different perspectives we might therefore expect to find that the schools with more positive health outcomes are characterised by stability of the student and staff body, good relationships between staff and students and a positive school ethos of shared norms.

Theory of human functioning and school organisation

We judged this theory as offering the most specific guidance about the mechanisms by which schools might determine health.⁶² It should be noted that this theory is not synonymous with the WHO guidance on HPS. It was produced by academic researchers not working with the WHO.

This theory asserts that a person's autonomy to make and enact good decisions is a necessary precondition for healthy behaviour. Informed by Nussbaum,¹⁰⁴ this theory outlines how young people have various needs that must be met and capacities which must be built in order to achieve such autonomy. Enabling young people to develop 'practical reasoning' and 'affiliation' is key because fulfilment of all other needs and capacities will require a person to be able to think and form relationships. 'Practical reasoning' involves an ability to understand and manage one's own feelings, perspectives and emotions, and appreciate that other people also have their own feelings, perspectives and emotions. Practical reasoning also involves considering different options when making a decision on how to behave, including thinking about one's own and others' perspectives, feeling and emotions. 'Affiliation' involves an ability to form relationships with others.

A school can enable its students to fulfil these capacities through what Bernstein¹⁰⁵ had previously called its 'instructional' and 'regulatory' orders. The instructional order is the way in which a school enables students to learn. It has traditionally involved developing students' practical reasoning through in-depth study of discrete academic subjects, but it can also refer to the development of life skills and emotional and social literacy. The regulatory order is the way in which a school aims to encourage norms of good behaviour and students' sense of belonging in the school community. Bernstein¹⁰⁵ argued that schools should aim to ensure their students are 'committed': engaged with and able to meet the challenges of the instructional order, and accepting the norms of, and feeling a sense of belonging to, the regulatory order. However, students can become 'alienated', 'detached' or 'estranged' (Table 2).

Informed by Bernstein,¹⁰⁵ Markham and Aveyard⁶² argue that alienated or detached students might instead seek alternative affiliation and self-development in other groups, such as anti-school peer groups, with consequences for health behaviours such as substance use, violence and teenage pregnancy. How students respond to school may depend partly on their social class and the extent to which school culture seems to connect with or contradict the culture that students experience in their families and communities. According

TABLE 2 Bernstein's categorisation of students based on responses to the 'instructional' and 'regulatory' orders

	Accept and meet challenges of the 'instructional' order	Reject or unable to meet challenges of the 'instructional' order
Accept values of the 'regulatory' order	Committed	Estranged
Do not accept values of the 'regulatory' order	Detached	Alienated

to Bernstein,¹⁰⁵ working-class students are more likely to become alienated, detached or estranged than middle-class students because they are more likely to feel that the school's culture does not resonate with their own culture and, therefore, that its instructional and regulatory orders are not aimed at meeting their own needs.

The theory of human functioning and school organisation suggests that schools will differ in how inclusive their culture is and the extent to which they enable different students to become committed. The extent to which schools are able to do this will depend on their modes of 'classification' and 'framing'.¹⁰⁵ 'Classification' refers to how the school as an institution and its curriculum are organised and, within this, how rigidly 'boundaries' are set. These boundaries can involve those between staff and students and those within the student body (e.g. through academic streaming, those learning different academic subjects/ studying for different qualifications). Some schools will reduce these boundaries and Bernstein¹⁰⁵ proposes that these schools will be more successful at building student commitment and promoting student autonomy and health. 'Framing' refers to the style in which staff communicate with and teach students, either rigidly, in which communication is teacher centred and teaching is didactic, or more flexibly framed, whereby communication is more equal and students are able to contribute to decisions about how learning proceeds.¹⁰⁶ This theory suggests that schools that maintain rigid social boundaries, between staff and students and/or among students, and which frame learning in teacher-centred rather than student-centred ways will fail to ensure that their students are committed, so that these students reject the values of the school and seek affiliation elsewhere, including with peer groups that embrace substance use and other risk behaviours.

Secondary theories that explain the mechanisms by which schools determine health

We identified a further 10 theories that could be used to suggest our secondary review hypotheses. Several of these theories did not meet our quality criterion of addressing a range of health outcomes but we judged them to be sufficiently useful in understanding school effects on ASB-related outcomes, hence we included them. The theory of triadic influence was included despite not meeting our quality criterion of simplicity because of its clarity and comprehensiveness, while the ecological model of co-ordinated school health programmes was included despite not meeting our quality criteria of simplicity, operationalisable constructs and clear pathways because, unlike all other theories, it attended to certain aspects of the school environment, such as school safety and opportunities for physical activity.

Flay's⁵⁶ *triadic theory of health behaviours* suggests that health behaviours are influenced by factors from three domains: intrapersonal factors (social competence and sense of self), socioenvironmental factors (behaviours of others and bonding to others) and the broader cultural environment (information and opportunities about behaviours and culture/religion). Each of these streams has distal and proximal elements moving from the social–personal nexus to expectancies and evaluations and to cognitions and affect, with dynamic inter-relationships between these. Attitudes, socially normative beliefs and self-efficacy determine decisions/intentions and behaviour. Although his theory offers a comprehensive framework for understanding how the influences on health behaviours inter-relate, it does not aim to offer a specific theory of how the school institution structures these factors. Nonetheless, it does suggest the hypothesis that schools may foster healthy behaviours by providing access to knowledge through health education, but also by reducing students' opportunities for engaging in risk and setting rules against risky behaviours, by

providing opportunities for bonding with staff and other students, and enabling the development of social skills through health and general education.

Eight theories outline why certain young people may engage in antisocial or risk-taking behaviours. *Social control theory*¹⁰⁷ suggests that individuals with a stake in a particular community will avoid committing acts considered deviant within that community. This might suggest the hypothesis that schools can reduce ASB by giving students some stake in their school community, perhaps by increasing their participation in decisions. The *integrated perspective on delinquent behaviour*¹⁰⁸ suggests that delinquency will be greater among individuals who have experienced a failure to participate in conventional social settings. This would suggest the hypothesis that schools can reduce ASB by ensuring that all students experience success in school activities. *Problem behaviour theory*¹⁰⁹ suggests that young people engage in behaviours such as drug use or risky sexual behaviour to cope with problems dealing with their wider system of conventional behaviour such as educational failure and low self-esteem. This would inform a hypothesis that schools could reduce ASB by ensuring that students' educational and social problems are addressed. *Strain theory*¹¹⁰ suggests that individuals may engage in ASB when they experience a strain between achieving what they regard as socially legitimate goals and their ability to achieve these through socially legitimate means. Thus, we might hypothesise that schools with lower rates of ASB are better at ensuring that students can achieve their broader goals through school activities.

None of the above theories considers the specific means by which schools may affect these mechanisms. Other theories go a little further towards suggesting what particular aspects of an institution might determine behaviour. For example, *deterrence theory*¹¹¹ would suggest the hypothesis that individuals will be deterred from behaviours if these are met with certain, severe and rapid punishments. Similarly, although the *theories of reasoned action*¹¹² and the *theory of planned behaviour*¹¹³ do not consider how the school environment is likely to influence health, they do suggest that behaviours that are the subject of clear sanctions within schools might be inhibited by encouraging students' acceptance of institutional norms and motivation to conform. This would suggest the hypothesis that schools with strict and strongly enforced codes against activities such as smoking, drinking and violence have lower rates of these outcomes. In contrast, *moral authority theory*¹¹⁴ argues that a prime aim of schools is to inculcate respect for the specific rules of the school as well as broader rules of social behaviour. However, this would not need to occur through strict enforcement and severe punishments, because acceptance of the rules can be internalised without recourse to such formal processes. This theory would suggest the hypothesis that schools with lower rates of ASB have a positive ethos and do not necessarily have strict rules, although the theory does not offer suggestions as to what system of organisation would be required to foster this positive ethos.

The *ecological model of co-ordinated school health programmes*,⁵⁷ although not offering a very deep understanding of how institutions affect health, does direct attention to particular aspects of schools that might promote health across multiple domains. This model would suggest a hypothesis that schools can foster health by promoting a supportive psychosocial environment and safe facilities, as well as opportunities for physical activity within the school.

Discussion

Summary of key findings

Twenty-four theories were identified in the in-depth theory synthesis. The theories most commonly cited in empirical reports were ecological systems theory, social control theory, social disorganisation theory, social learning theory, the theory of human functioning and school organisation and social cognitive theory. We considered several criteria to decide which theories to use to inform our primary and secondary review hypotheses.

Table 3 indicates the testable hypotheses that the primary and secondary theories we identified might suggest.

TABLE 3 Testable hypotheses of included studies

Theory	Hypothesis
Primary theories	
Social capital theory	Schools foster health by having a stable student and staff body, good relationships between staff and students and a positive school ethos of stable, shared norms
Social development model	Schools reduce ASB by providing opportunities for students to participate fully in learning and community life and develop the skills necessary for such participation and ultimately enabling students to gain recognition
Theory of human functioning and school organisation	Schools foster student autonomy and health by reducing social boundaries between staff and students and among students, and ensuring student-centred framing of learning, management and other school systems
Secondary theories	
Deterrence theory	Schools reduce ASB by setting certain, severe and rapid punishments
Theory of reasoned action	
Theory of planned behaviour	
Ecological model of co-ordinated school health programmes	Schools foster health by promoting a supportive psychosocial environment, good safety facilities and opportunities/requirements for physical activity within the school
Integrated perspective on delinquent behaviour	Schools reduce ASB by ensuring that all students experience success in school activities
Moral authority theory	Schools reduce ASB by inculcating respect and not necessarily setting severe punishments
Problem behaviour theory	Schools reduce ASB by ensuring that students' educational and social problems are addressed
Social control theory	Schools reduce ASB by giving students some stake in the school community, perhaps by increasing student participation in decisions
Strain theory	Schools reduce ASB by ensuring that students can achieve their broader goals through school activities
Theory of triadic influence	Schools foster health by providing health education, reducing students' opportunities for engaging in risk, setting rules/norms against risky behaviours, enabling bonding between staff and students and providing good general education

Strengths and limitations

Our initial summary of theoretical literature in stage 1 was relatively unsystematic: we noted theories that recurred in the first half of our coding for the evidence map but not in the second. The preliminary summary identified only five of the theories that were identified in the in-depth synthesis. Nonetheless, this provided us with some insights into the range of theories informing the empirical studies. Along with the evidence map it enabled us to have a lively discussion with stakeholders about which types of evidence it would be most interesting and useful to review in depth in stage 2. We cannot rule out the possibility, however, that a more comprehensive summary of theory at this stage would have led to different priorities.

Our summary and assessment of theories in stage 2 was systematic, using a tool of our own devising. The judgements we made were to some extent subjective, for example in determining whether or not a theory was simple and had constructs that could be operationalised in empirical research. We used these multiple criteria to form a judgement about which theories to draw on to define our primary and secondary review hypotheses. We did not simply require that a theory meet every criterion in order to be

considered primary because our judgements were necessarily more subtle than this. For example, the theory of human functioning and school organisation did not meet our criterion of simplicity; however, we judged that this theory was very strong regarding our other criteria and so included it. As a further example, several of our theories did not meet our quality criterion of addressing a range of health outcomes but we judged them to be sufficiently useful in understanding school effects on ASB-related outcomes and so we opted to use them to inform secondary review hypotheses. We think this balance between using clear criteria and making overall judgements is acceptable and appropriate given that these concerned the development rather than the testing of hypotheses.

Despite its subjectivity, this process was useful in determining which theories could most usefully provide hypotheses to assess against the empirical reviews. These theories enable us to develop hypotheses about how school environment interventions and school-level exposures might affect health, but did not enable us to focus on specific pre-hypothesised outcomes.

The theories themselves were biased towards those focusing on ASB, with six of our secondary theories but no primary theories having this focus. However, this reflects the theories that were used in empirical studies of the health effects of schools and school environment interventions and is an interesting finding of our review.

Chapter 7 Research question 2: outcome evaluations

Research question

What are the effects of school environment interventions (interventions aiming to promote health by modifying how schools are organised and managed, or how they teach, provide pastoral care to and discipline students, and/or the school physical environment) that do not include health education or health services as intervention components and which are evaluated using prospective experimental and quasi-experimental designs, compared with standard school practices, on student health (physical and emotional/mental health and well-being; intermediate health measures such as health behaviours, BMI and teenage pregnancy; and health promotion outcomes such as health-related knowledge and attitudes) and health inequalities among school students aged 4–18 years? What are their direct and indirect costs?

Methods

Inclusion and exclusion criteria

A total of 524 references were identified in the evidence and theory map as relevant to RQ2. We included experimental and quasi-experimental evaluations of the effects on health or well-being outcomes in students (aged 4–18 years) of school environment interventions that addressed how schools were led and managed, how they teach, support or discipline students, or schools' physical environment. Two reviewers (CB and HW) independently double-sifted these references based on title and abstract only and on full reports where necessary using the exclusion criteria in *Table 4*. It should be noted that all references coded as outcome evaluations were screened for inclusion because of the potential limitations of the accuracy of the coding in the stage 1 evidence map. Screening was not hierarchical or mutually exclusive. In other words, references may have been excluded based on multiple criteria.

Quality assessment

All included reports were quality assessed using the following criteria:

- whether or not the impact of the intervention was reported for all outcomes
- whether or not allocation to the intervention and comparison conditions was random
- whether allocation was non-randomised but involved matching or adjustment for potential confounders
- whether or not there were equivalent study groups at baseline
- whether or not attrition overall is $\leq 30\%$
- whether or not attrition rates differed by $\leq 10\%$ between groups
- whether or not analysis accounted for cluster.

These criteria used for assessing methodological quality were adapted from those used in EPPI-Centre health promotion reviews.¹¹⁵ These allowed us to assess which studies were most likely to provide us with estimates of the effects of school environment interventions minimally biased by confounding, selection and information bias and random error. Based on the results of these criteria, reviewers judged the study overall as well conducted or less well conducted.

These quality assessment criteria were piloted on a random sample of two reports by two reviewers (CB and HW) before being applied by one reviewer (HW) and checked by another (CB), with any differences being settled by discussion without recourse to a third reviewer. These criteria were not used to include or exclude

TABLE 4 Exclusion criteria for intervention outcome evaluations (in-depth synthesis)

Exclusion criterion	Guidance
Exclude 1: not an evaluation	Exclude if study is not an evaluation study
Exclude 2: is a process evaluation only	Exclude if the study is a process evaluation
Exclude 3: based on intervention	Exclude if the study intervention does not address how schools are led and managed, or how they teach, support and discipline students, and/or the school physical environment (e.g. intervention merely involves extracurricular activities, catering, PE or active transport). Exclude if the intervention includes curriculum and community/parent components alongside school environment components With intervention outcome evaluation studies we were interested in intervention studies in which the intervention aimed to modify student–student or staff–student relationships, as long as they did this by addressing the school environment and not merely through health education
Exclude 4: not a cluster RCT+ non-randomised prospective	Exclude the report if it is an outcome evaluation, but does not involve (a) a cluster RCT or (b) a non-randomised prospective cluster comparison design
Exclude 5: (Cochrane) curriculum/parent	Exclude if the intervention includes curriculum and community/parent components alongside school environment components [these reports were passed on to the Cochrane group (Bristol/Cardiff colleagues)]
Exclude 6: other	Exclude report for other reason(s) (specify) (could not be located; non-English duplicate of already coded reference; no relevant health outcomes; teacher not student health; topic)
Include	Study meets the criteria for in-depth synthesis

reports from the in-depth review in stage 2, but they were used to determine what weight we gave study findings in our narrative synthesis.

Data extraction

For the included RCT and quasi-experimental study reports we extracted data related to intervention (description, theory, length, providers internal/external to school), comparison group, study RQs/hypotheses, study site and population, sampling, schools and individuals per arm, data collection methods, analysis methods, results and authors' conclusions. We also extracted data on adjustment for clustering, confounders and methods to control these, waves of follow-up and attrition rates overall and by study arm, outcome measures, effect size estimates (overall and by subgroup) and measures of confidence/significance, intervention costs and indirect resource use and economic measures of cost-effectiveness. Data extraction tools were piloted on a random sample of two reports by two reviewers (CB and HW) before being applied by one reviewer (HW) and checked by another (CB), with any differences being settled by discussion.

Synthesis

We assessed the potential for statistical meta-analysis of intervention outcome evaluations by noting which reports focused on similar combinations of school-level interventions and outcomes. We set a threshold of a minimum of three such reports being required to consider meta-analysis because meta-analysis of only two studies is unlikely to produce stable effect estimates. Valentine *et al.*¹¹⁶ suggest that meta-analysis of two studies will be informative only when studies are direct replications. With very few studies, parameter estimation such as point estimates will likely be poor, rendering conclusions uncertain. We found no homogeneity of interventions and outcomes whatsoever for our intervention studies. Therefore, we subjected our quantitative study reports of intervention effects to narrative synthesis. In conducting our narrative synthesis of intervention outcome reports we included all reports in the narrative but made clear where studies were subject to major methodological limitations.

There were three aspects to our narrative synthesis of evidence from intervention outcome evaluations. First, we prepared a narrative summary of the findings from each study. We summarised reported effects in the

context of study quality, setting and population. This is reported in the results section of this chapter. Second, we developed a narrative overview of findings from across our outcome evaluations. We noted where studies of similar interventions or influences reported convergent or divergent findings and considered the reasons for this in terms of design and context, rather than merely weighing the overall frequency of differing findings. This is reported in the discussion section of this chapter. Finally, we examined the implications of the findings from outcome evaluations alongside those from other forms of evidence reviewed for our review hypotheses. Because our review hypotheses were worded generally rather than specifically in terms of precise constructs, and because empirical studies had not been designed to test our review hypotheses, we did not test hypotheses in a statistical sense but in a more general narrative sense, to the extent to which these findings supported or undermined the review hypotheses. This is reported in *Chapter 11*.

Overview of included reports

Flow of literature

Of the 524 references identified in the evidence map, 514 were excluded: 83 were not an outcome evaluation, 47 were process evaluations only, 251 did not meet our criterion of a school environment intervention addressing school organisation or management, teaching, pastoral care or discipline, or physical environment (e.g. they included health education curricula, or addressed school catering, PE, etc.), 134 were not a cluster RCT or non-randomised prospective study, 85 were related to curriculum or parents and nine were excluded for other reasons (one could not be located, one was a duplicate, two had no relevant health outcomes, one was about teacher health and four were not related to our topic) (all of the study reports were written in English) (*Figure 8*).

Two RCT reports relating to the same study were suggested by our Cochrane colleagues,^{54,55} one quasi-experimental study was suggested by CB⁵⁹ and three references were located by reference sifting.^{58,117,118} In total, 10 studies (16 reports) from the evidence map were included in the in-depth synthesis (of which six studies were RCTs and four were non-randomised outcome evaluations).

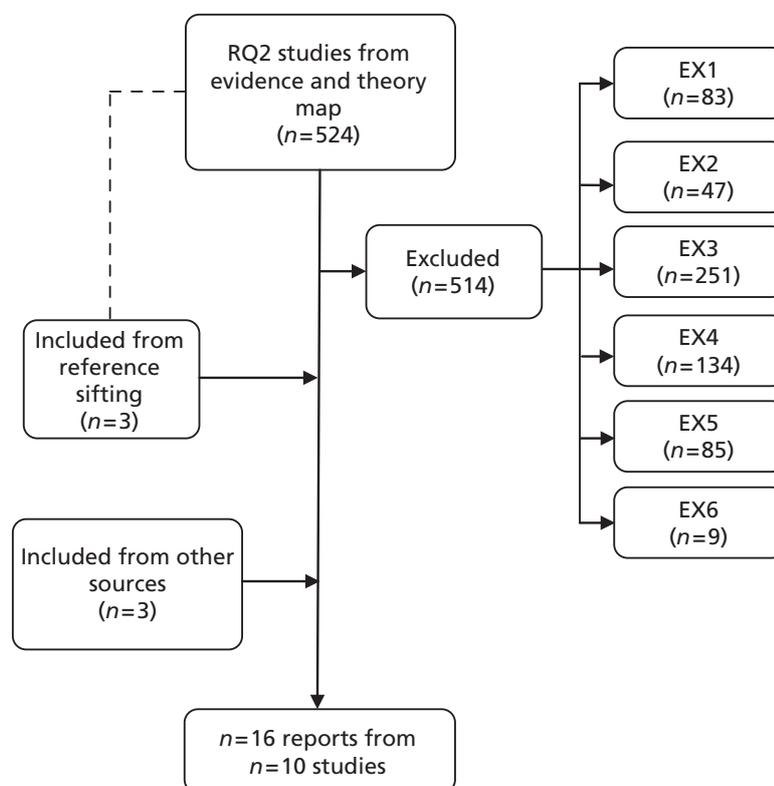


FIGURE 8 Flow of literature: outcome evaluation synthesis (stage 2: in-depth synthesis).

Quality assessment

The methodological quality of the studies was judged against the seven quality criteria. All 10 studies^{44,54,55,58,63,119–122} were deemed suitable to include in the narrative synthesis, although the quality of the studies varied considerably. Six studies employed a RCT design, although this was compromised in Flannery *et al.*¹¹⁹ by the replacement of a dropout school with a non-randomised substitute. Four studies involved a quasi-experimental prospective comparison group design. Those reported by Sallis *et al.*,¹²² Dzewaltowski *et al.*⁶³ and Flay *et al.*⁴⁴ were the strongest studies in terms of design, sample size and adjusting for clustering in the analysis.

All of the RCT studies except for Flay *et al.*⁴⁴ involved arms that compared school environment interventions with normal practice. Fonagy *et al.*¹²⁰ also involved a third arm that was not relevant to our RQs. Flay *et al.*⁴⁴ included three arms: school environment intervention plus curriculum, curriculum only and normal practice. We included this trial because comparison of the first two arms enabled us to examine the effects of school environment over and above the effects of the curriculum element and so this was analogous to comparing school environment only with normal practice.

Study characteristics

Ten RCT reports^{44,54,55,58,63,119–123} examining six interventions and six reports^{60,117,118,124–126} from quasi-experimental outcome evaluations examining four interventions were included in the in-depth review. They were published between 1988 and 2010 and evaluated multicomponent interventions delivered in elementary and middle schools in the USA and elementary (primary) and secondary schools in England. The interventions fell into three categories: five aimed to encourage staff and students to develop school climates characterised by a stronger sense of community and/or better interpersonal relations to reduce aggression and other risk behaviours;^{44,54,55,58,60,119–121,123} two encouraged staff and students to advocate for school environments promoting healthier eating and physical activities;^{63,122} and three involved improving school playgrounds.^{117,118,124–126} All of the interventions were universal in addressing all students. None of these studies reported any data on cost-effectiveness.

Intervention details of randomised controlled trials

Eight RCT reports,^{44,54,55,58,119–121,123} from four studies, evaluated interventions that encouraged staff and students to develop school climates characterised by a stronger sense of community and/or better interpersonal relations. The way in which the interventions map against the domains specified in our review hypotheses is summarised in *Chapter 11* (see *Table 8*).

Battistich *et al.*,⁵⁵ Battistich *et al.*,⁵⁴ Solomon *et al.*⁵⁸ and Battistich¹²³ report the effects of the Child Development Project (CDP). The project was delivered over 5 school years to students entering elementary schools in northern California in 1982. The project aimed to promote prosocial behaviour by providing children with experiences that engender a sense of community and a climate of mutual respect and concern in the classroom and school. Teachers were trained in establishing a caring classroom environment in which children were enabled to learn about others' needs, feelings and perspectives, collaborate with one another, consider issues of fairness, kindness and social responsibility, and participate in decisions about their activities and their classrooms. The main input to this 5-year programme was training (curriculum, workshops, coaching) for teachers. The study reports do not provide sufficient details about the duration of training inputs or the skill level of trainers to inform estimates of intervention costs. The study reports also do not report the opportunity and other costs incurred by schools as a result. The trial involved three schools per arm, which reduced to two per arm at later follow-ups. Follow-up rates for students were low and incompletely reported and analysis did not account for clustering at school level.

Flay *et al.*⁴⁴ report the effects of the Aban Aya Youth Project (AAYP). The work was conducted from 1994 to 1998 in elementary and middle schools in Chicago, IL, USA, predominantly attended by African American students. This well-conducted and -reported RCT compared the school/community intervention (SCI) arm, which included school environment and a social skills curriculum, with the social development curriculum

(SDC)-only arm as well as a no-intervention comparison arm, and thus enabled the effects of the school environment component to be distinguished from the effects of the curriculum. Six schools were involved in each arm. Researchers followed students from grade five through to grade eight. Although follow-up rates of original participants were low, the analysis focused on repeat cross-sectional summaries of school-level outcomes. The SCI involved schools forming a school task force consisting of school personnel, students, parents, community advocates and project staff to implement the programme components, propose changes in school policy, develop other school–community liaisons supportive of school-based efforts, and solicit community organisations to conduct activities to support SCI efforts, with the aim of ‘rebuilding the village’ and creating a ‘sense of ownership’ of the school. This was supported by staff training enabling staff to integrate prosocial skills into the school environment and enhance classroom management and interactive teaching skills. The trial report provides detail on the intervention, but not enough information on activities to inform costing. For example, it is not reported how often the task force met. Similarly, although it is reported that there was staff in-service training, the materials and methods are not discussed in detail. Although some sample schools’ actions, such as reviews of school policy and school-wide fairs, are discussed, it is not possible to develop a sense of the overall amount of activity or its resource consequences.

Fonagy *et al.*¹²⁰ report the effects of the Creating a Peaceful School Learning Environment (CAPSLE) intervention. The work was conducted in elementary schools in Kansas (date not specified), USA. This RCT also involved three arms in which CAPSLE was compared with both school psychiatric consultation, a psychiatric intervention targeting at-risk students, and a no-intervention comparison arm, following all third to fifth graders for 2 years of active intervention followed by 1 further year of follow-up. Our review examines differences only between the CAPSLE and no-treatment arms. CAPSLE aimed to enable all staff and students to interpret their own and others’ behaviour and thereby counteract bullying of others. Teachers received group training and students engaged in peer mediation programmes, receiving self-defence training (this aiming to facilitate the intervention rather than being part of the conventional health curriculum). Although the trial report provides quite detailed information on the initial and refresher training provided to staff and students, it is not clear who provides this or how much time was involved. It is reported that biweekly supervision meetings were provided but it is less clear how much time is spent on ongoing consultation or who provides this. It is also not clear what are the ramifications of these inputs in terms of school activities and opportunity and other costs so that it is not possible to estimate accurately what resources the intervention requires. Three schools were allocated to each arm. Loss to follow-up was high and differential by arm, and analysis did not account for school clustering.

Flannery *et al.*¹¹⁹ and Krug *et al.*¹²¹ report the effects of PeaceBuilders, a universal violence prevention programme that attempts to alter the climate of a school by teaching students and staff rules and activities aimed at improving child social competence and reducing aggressive behaviour, weaving these into everyday routine rather than presenting them as a separate curriculum. The work was conducted in elementary schools in the vicinity of Tucson, AZ, USA. Flannery *et al.*¹¹⁹ report quite detailed information on the initial inputs provided to schools to deliver this intervention in terms of provider and timing but subsequent inputs are described only vaguely in terms of this being on an as-needed basis; involving specific sessions on issues identified by staff in each school; involving periodic group meetings to discuss successes and challenges; and involving occasional 1-day meetings. There is also no reporting of the time commitments and opportunity and other costs of the intervention accruing to schools in the course of their delivering the intervention. Four schools were randomly allocated per arm and students in kindergarten to fifth grade were followed up for 1 year within the randomised trial, with high attrition. New kindergarten students were not included in the follow-up sample and those in fifth grade at baseline were not followed up.

Reports from two RCTs examine an intervention that enabled staff and students to advocate for school environments promoting healthier eating and physical activities. These interventions were included not because of their outputs in terms of modifying school nutrition and physical activity environments but because of their processes of involving students working alongside staff in making decisions. These interventions were also universal, addressing all students.

Dzewaltowski *et al.*⁶³ report the effects of the Healthy Youth Places (HYP) intervention, which aimed to develop the 'personal and proxy agency' of adult leaders and youth to ensure that schools promote healthy eating and physical activity. The work occurred in US middle schools during an unspecified time period. The intervention focused on engaging students in changing the environment. This was to occur through school site co-ordinators facilitating youth-led school advocacy groups, known as 'change teams', which were the hub of intervention activities at the school, also involving school staff. Local capacity to participate in intervention activities was enabled through training for staff and students on leadership skills. The intervention outcome evaluation report provides some intervention detail but not sufficient to cost the intervention. It is reported that expert staff delivered continuous group staff training to paid school site co-ordinators from the eight intervention schools but it is not reported how much training in total was provided, how expert the trainers were or whether paid co-ordinators were employed full- or part-time. The facilitation of the youth-led school advocacy groups by the site co-ordinators was supported by a training programme for youth in school change, but insufficient detail is provided to estimate costs. It is also not reported what are the opportunity and other costs to schools, nor are sufficient details given on the actions taken to provide an accurate estimate of this. Eight schools were allocated per arm with students followed from the end of sixth grade through to eighth grade, with reasonably good rates of follow-up.

Sallis *et al.*¹²² report the effects of the Middle-School Physical Activity and Nutrition (M-SPAN) intervention, which aimed to increase physical activity and reduce dietary fat intake at school. The work proceeded at middle schools in San Diego, CA, USA, from 1997 to 1999. School teaching and catering staff worked alongside students to develop action plans involving rewriting school policies and generating environmental change. Student health committees worked alongside a teacher to support implementation of activities. Parental education was delivered through existing school communication channels. This intervention is included as it focused on engaging students in changing the environment. The trial report provides detailed information on the intervention but not enough to inform cost estimates. For example, although it is reported who attended meetings between the trial team and schools and how often these meetings occurred, the same information is not provided for schools' internal planning meetings. The opportunity and other costs to schools of actions such as closing school stores at lunchtime were also not provided. It is also not clear how often student health committees met or what were the costs of the activities in which these engaged. The trial reports the amount of funding provided to schools as an incentive to participate, and for kitchen and PE equipment, but it is not clear, and is unlikely to be the case, that this met the actual costs of the intervention to schools. Within this well-conducted trial, 12 schools were allocated per arm and, as with Flay *et al.*,⁴⁴ analysis focused on repeat cross-sectional data with observations and student/parent surveys being conducted on separate random samples at baseline and follow-up.

Intervention details of quasi-experimental studies

Six reports of quasi-experimental outcome evaluations evaluated four interventions.

Bonell *et al.*⁶⁰ report on the Healthy School Ethos (HSE) intervention, which encouraged staff and students to develop school climates characterised by a stronger sense of community and better social skills. The intervention involved each intervention secondary school (serving students aged 11–16 years in south-east England) instituting an action team to plan and deliver actions to improve students' relationships with teachers and other students, sense of security, social support, self-regard and engagement, to reduce substance use. The action team comprised staff, students from multiple year groups, and parents. Actions were either mandatory (e.g. revising school policies and rules) or locally determined (e.g. developing 'safe spaces' for younger students, training older students as peer mediators, resolving disputes and preventing bullying). The intervention included a facilitator with accompanying manual, staff training, funding and a survey of student needs. There was no curriculum component. The HSE study report gives detailed information about intervention inputs and some processes. However, the intervention enabled a multitude of locally determined actions of varying degrees of likely intensity (ranging from overhauling a school rewards policy to implementing new forms of peer mediation) in each school, which are not described in sufficient detail to be costed but are likely to form a significant proportion of the costs. Although conceived as a small pilot matched-pair cluster RCT involving two intervention and two comparison schools, the dropout of one

school allocated to the intervention arm prior to delivery required a switch to a non-randomised matched design involving 1 year of follow-up tracking year 7 students (age 11 years). Rates of follow-up were high and similar between arms.

Ridgers *et al.*^{124,126} report on an intervention aiming to improve playground physical structures. This was delivered in elementary schools (grades K–4) in north-west England. Each school received £20,000 (total £300,000) to redesign the playground environment into three specific zonal coloured areas: (1) a red zone, (2) a blue zone and (3) a yellow zone. The red zone was a designated sports area with physical markings for sports (soccer, tennis, basketball). The blue zone served as a fitness and skills area (multiactivity) and the yellow zone served as a 'chill-out'/quiet play area. The intervention schools received the following physical structures: soccer goal posts, basketball hoops, fencing around the red sports area and seating in the yellow quiet area. None of the supervisors received training in the promotion of playground physical activity. The only costs not reported are the staff time of the schools for managing the process of improving the playground areas. The study involved a matched comparison group prospective design with students followed up at 6 weeks, 6 months and 12 months. There is some inconsistency in the reporting of how many schools participated, with Ridgers *et al.*^{124,125} reporting 15 intervention schools and 11 control schools at the 6-week and 6-month follow-ups and another study by Ridgers *et al.*¹²⁶ reporting 10 control schools at the 6-month follow-up and seven schools at the 12-month follow-up, with no explanation of these disparities.

Stratton¹¹⁵ and Stratton and Mullan¹¹⁸ evaluated separate interventions to improve school playground surface markings in primary schools in north-west England. These involved painting bright fluorescent colours on the playground surface. Stratton and Mullan¹¹⁸ report that this cost £800 per school. Both studies involved non-randomised prospective comparison groups. The study by Stratton¹¹⁷ involved only one intervention and one comparison school, while Stratton and Mullan's study¹¹⁸ involved four intervention and four comparison schools. Stratton¹¹⁷ matched the schools on area deprivation and playground area as well as participating children's age, size and stature. Stratton and Mullan¹¹⁸ matched schools on playground dimensions and area deprivation. Both studies experienced poor rates of follow-up, which was differential by arm, and did not adjust for clustering in their analysis.

Results

Narrative summary of findings from randomised controlled trials

Interventions that encouraged staff and students to develop school climates characterised by a stronger sense of community and better relationships appeared to bring significant benefits for some but not all health outcomes. However, most of these evaluations had important limitations, and the strongest evaluation in this category reported the fewest significant benefits.

Battistich *et al.*⁵⁵ report that the CDP intervention had positive effects on three of the four measures of cognitive social problem-solving in conflict situations among grade four US elementary school students when followed up at grade four. There was an interaction between arm and grade indicating that the intervention effects increased between kindergarten and grade four, suggesting that benefits were cumulative. Battistich *et al.*⁵⁴ report no programme effects for self-esteem at grade four but benefits for loneliness/social dissatisfaction as well as social anxiety at grade six. Solomon *et al.*⁵⁸ also report a positive effect for teacher-reported students' solving of minor interpersonal problems, drawing on cumulative data from kindergarten through to fourth grade. Battistich¹²³ reports intervention effects reducing student self-reported loneliness/social dissatisfaction and social anxiety at grade six but not self-esteem at grade four. However, this evaluation had a number of weaknesses. Although the evaluation involved a RCT design, baseline data were not collected; it is therefore not possible to assess whether or not study groups were equivalent at baseline and so it is impossible to assess the risk of confounding. Furthermore, attrition was high overall and differed markedly between groups so that selection bias is a risk. Finally, no account is taken in the analysis of the clustering of data within schools. Therefore, the results of this study should be interpreted with caution.

Flay *et al.*,⁴⁴ reporting on the AAYP in Chicago elementary and middle schools, compared the arm of their RCT involving a school environment component plus curriculum with the curriculum-only arm. Among boys there were non-significant trends in the direction of benefit for all of the health outcomes (violence, recent sexual intercourse, condom use) at the 3-year follow-up in grade eight. However, only for two non-health outcomes (school delinquency and an overall aggregate measure combining the health outcomes with school delinquency and provoking behaviours) were the differences statistically significant. Among girls there was a significant increase in self-reported condom use ($p=0.03$), but no effects for other measures. This was a well-conducted RCT with little risk of confounding or bias.

Fonagy *et al.*¹²⁰ report that, compared with no intervention, the CAPSLE intervention reduced the developmental trend of increasing peer-reported victimisation ($p<0.01$), peer-reported aggression ($p<0.05$), self-reported aggression ($p<0.05$) and peer-reported aggressive bystanding ($p<0.05$) and increased peer-reported helpful bystanding ($p<0.001$) among third to fifth graders in US elementary schools at the 2-year follow-up. CAPSLE also reduced a decline in self-reported empathy ($p<0.01$). These benefits were not maintained at the 3-year follow-up other than for peer-reported helpful bystanding ($p<0.05$). Results for self-reported victimisation and self-reported beliefs in the legitimacy of aggression did not suggest intervention benefits at 2 or 3 years' follow-up. This evaluation involved a RCT design but with high attrition that differed between groups so that selection bias is a risk. No account is taken in the analysis of the clustering of data within schools. Therefore, the results of this study should be interpreted with some caution.

Regarding the PeaceBuilders intervention delivered in US elementary schools, Flannery *et al.*¹¹⁹ report that, at the 1-year follow-up, compared with students in control schools, teachers rated third to fifth grade but not K–2 grade students in intervention schools as significantly lower in log-aggression, and there were no intervention effects on child self-reported aggression. The intervention effect for teacher-rated aggression was larger for students with higher aggression scores at baseline. However, subgroup effects by socioeconomic status (SES), ethnicity or sex are not reported and a test for interaction with baseline aggression was not performed. For the same intervention, Krug *et al.*¹²¹ report that, overall, injury-related and fighting-related as well as non-fighting-related rates of visits to the school nurse decreased significantly more in intervention than in comparison schools at the 1-year follow-up while visits from injuries possibly related to fighting did not differ. This evaluation involved a RCT design. One comparison school dropped out at baseline and was replaced with another non-randomised school suggesting that, along with high attrition, which differed between groups, there is some risk of selection bias. The results of this study should therefore be interpreted with some caution.

Both interventions that aimed to address physical activity and healthy eating appeared to affect the former but not the latter. Dziewaltowski *et al.*⁶³ report that, immediately after the HYP intervention, US middle schools did not change their fruit and vegetable consumption but did significantly change their levels of physical activity compared with control schools. From sixth to eighth grade the intervention was associated with increases in some but not all measures of vigorous physical activity [vigorous physical activity (VPA), $p=0.03$] and moderate/vigorous physical activity [moderate/vigorous physical activity (MVPA), $p=0.005$]. There were significant increases in students' efficacy to ensure that schools were creating supportive environments for physical activity as well as their self-efficacy to be physically active, but not in their efficacy with regard to parent or peers creating supportive environments for physical activity. The intervention did not have effects on any outcomes relating to healthy eating other than their own self-efficacy to eat five to seven portions of fruit and vegetables per day and group norms supportive of fruit and vegetable consumption. Sex, race or SES did not interact with the intervention effects over time. Inclusion of a measure of student proxy efficacy in the model for the effect of the intervention on MVPA and VPA suggested that this was a mediator for intervention effects ($p=0.05$ and 0.03 for MVPA and VPA respectively). This evaluation involved a RCT design but with high attrition that differed between groups so that selection bias is a risk. The results of this study should therefore be interpreted with some caution.

In US middle schools (student grades not reported), Sallis *et al.*¹²² report that there was a significant intervention effect of the M-SPAN intervention at the 1-year follow-up for observed physical activity for the

total group ($p < 0.009$) and for boys ($p < 0.001$), but not for girls ($p < 0.40$). The intervention was not effective for consumption of total fat ($p < 0.91$) or saturated fat ($p < 0.79$) at 2 years' follow-up. The intervention did not have a significant effect on self-reported physical activity or sedentary behaviours. Similarly, there was no intervention effect on fatty foods consumed or parental fat avoidance. There was a significant reduction in BMI among intervention boys compared with control boys ($p = 0.044$), but there was no effect for girls. This was a well-conducted RCT; however, the reported differences in effect were not subject to a test for interaction and so should be interpreted cautiously.

Narrative summary of findings from quasi-experimental studies

In English secondary schools, Bonell *et al.*⁶⁰ report that at the 1-year follow-up there were significantly higher reported rates of feeling safe at school among year 7 students in intervention schools [adjusted odds ratio (OR) 2.89; 95% confidence interval (CI) 1.63 to 5.15; $p < 0.001$] as well as differences that were of borderline significance (adjusted $p < 0.1$) in self-reported measures of teasing of others, hurting of others and been in a fight. There were no significant differences regarding self-reported measures of being teased or threatened weekly/hurt ever in this school, belief that will try illegal drugs or smoke a cigarette in the future or belief that will get drunk before age 16 years. This was a small pilot study involving a non-random design and so its findings should be interpreted with caution.

For their playground physical structure intervention delivered in English elementary schools, Ridgers *et al.*¹²⁵ report that, at the 6-week post-intervention follow-up, there were no significant differences between students in intervention and comparison schools in heart rate telemetry-derived and accelerometer-derived measures of physical activity. However, another study by Ridgers *et al.*¹²⁴ reports that at the 6-month follow-up the rates of four outcomes were significantly higher among students at intervention schools after adjustment for possible confounders: MVPA by heart rate (beta coefficient=4.03; $p = 0.042$); VPA by heart rate (beta coefficient=2.43; $p = 0.045$); MVPA by accelerometry (beta coefficient=4.53; $p = 0.025$); and VPA by accelerometry (beta coefficient=2.32; $p = 0.005$). At the 12-month follow-up, after adjustment, there were significant differences for MVPA by accelerometry (risk difference=1.2; 95% CI 2.8 to 5.2) and VPA by heart rate (risk difference=0.9; 95% CI 1.6 to 3.5). There were no significant differences for MVPA by heart rate, vigorous activity by heart rate, VPA by accelerometry measured at morning recess and MVPA by heart rate and VPA by heart rate measured at lunch recess. Ridgers *et al.*¹²⁵ report that, at the 6-week post-intervention follow-up for MVPA, there was a significant interaction between the intervention and age ($p = 0.01$), with intervention effects stronger for younger children, and an interaction of borderline significance for recess duration ($p = 0.07$), with effects stronger for longer recess period, but no interaction for baseline MVPA, sex or BMI. For VPA, there was a significant interaction between the intervention and age ($p = 0.09$), with intervention effects stronger for younger children, but no interaction for recess duration, baseline VPA, sex or BMI. Ridgers *et al.*¹²⁴ report that a positive interaction occurred between the intervention and recess duration for both MVPA and VPA ($p < 0.05$). Inverse interaction terms occurred between the intervention and baseline heart rate-measured MVPA and VPA ($p < 0.05$ and $p < 0.1$ respectively), indicating that the impact was stronger for children who were less active at baseline. An inverse interaction between the intervention and age occurred for MVPA ($p < 0.05$), suggesting that the impact was stronger for younger children. A positive interaction occurred between the intervention and time for VPA ($p < 0.05$), suggesting that the impact strengthened across time. Ridgers *et al.*¹²⁶ report significant interactions between the intervention and time for MVPA measured by heart rate and VPA measured by heart rate at morning recess, and MVPA measured by heart rate, VPA measured by heart rate and MVPA measured by accelerometry at lunch recess. These indicate that physical activity levels were lower at the 12-month follow-up than at the 6-month follow-up, with the greatest intervention effects observable at 6 months. Findings from this study should be interpreted with caution because of its non-random design and unclear reporting of participant follow-up rates.

Stratton¹¹⁷ reports no significant effects on VPA, MVPA or mean heart rate in the intervention school students compared with the comparison school students immediately after the intervention to provide playground surface markings. Comparing students in the four schools that received new playground markings with matched students in the four schools that did not receive this intervention, Stratton and Mullan¹¹⁸ report

significant effects immediately after the intervention on VPA ($F_{1,204}=4.05$; $p<0.03$) and MVPA ($F_{1,204}=13.7$; $p<0.01$). Findings from both of these evaluations should be interpreted with caution because of their non-random designs, variable rates of attrition between arms and failure to adjust for school clustering of data.

Discussion

Narrative overview of findings

Six RCTs (10 reports) and four quasi-experimental evaluation studies (six reports) were included, although the quality of these evaluations was generally quite poor, particularly in terms of attrition and adjustment for clustering in the analysis and small sample sizes. These intervention reports addressed most aspects of our second RQ, for example the AAYP and HSE projects modified the way that schools were managed and addressed pastoral care; CDP changed the way that classroom teaching occurred; CAPSLE and PeaceBuilders modified school discipline; and the playground interventions modified school physical environments. The study reports examined measures of health (such as injuries and social anxiety), health behaviours (such as aggression and condom use) and 'health promotion' outcomes (such as attitudes to aggression); however, they provided few data relevant to health inequalities or costs.

Nine papers reporting on five evaluations (all but one report/study involving a RCT) examined interventions that encouraged staff and students to build a stronger sense of community and/or better interpersonal relations.^{44,55,63,119,120} Such studies have been conducted in a range of school settings (elementary, middle and secondary/high schools), although all except the quasi-experimental evaluation of the HSE intervention (UK) were conducted in the USA. Studies reported a range of significant benefits for some but not all outcomes. Among US elementary school students, by grade four the CDP intervention was reported as having positive effects on three of the four measures of cognitive social problem-solving in conflict situations, as well as teacher-reported students' solving of minor interpersonal problems, but no effects for self-esteem. By grade 6 it had benefits for loneliness/social dissatisfaction and social anxiety, and reduced student self-reported loneliness/social dissatisfaction and social anxiety.^{54,55,58} The AAYP intervention was delivered in Chicago elementary and middle schools.⁴⁴ Compared with schools in the curriculum-only arm, schools receiving the school environment component plus curriculum experienced non-significant beneficial trends for violence, recent sexual intercourse and condom use among boys at 3 years' follow-up in grade eight. Only for two non-health outcomes (school delinquency and an overall aggregate measure combining the health outcomes with school delinquency and provoking behaviours) were differences statistically significant. Among girls there was a significant increase in self-reported condom use, but no other effects. The CAPSLE intervention reduced peer-reported victimisation and aggression, self-reported aggression and empathy, and peer-reported aggressive bystanding, and increased peer-reported helpful bystanding among third to fifth graders in US elementary schools at 2 years' follow-up.¹²⁰ CAPSLE also reduced a decline in self-reported empathy. These benefits were not maintained at the 3-year follow-up other than for peer-reported helpful bystanding. There were no effects for self-reported victimisation and self-reported beliefs in the legitimacy of aggression. The PeaceBuilders intervention delivered in US elementary schools had significant effects on teacher-rated aggression among third to fifth graders but not grade K–2 students, but there were no intervention effects on child self-reported aggression.^{119,121} Effects for teacher-rated aggression were larger for students with higher baseline aggression although no test for interaction was performed. At the 1-year follow-up there were intervention effects for injury-related and fighting-related visits to the school nurse. The HSE intervention was associated with significantly higher reported rates of feeling safe at school post intervention among year 7 students, and there were differences of borderline significance in self-reported teasing of others, hurting of others or been in a fight. There were no significant differences regarding self-reported measures of being teased or threatened weekly/hurt ever in this school, belief that will try illegal drugs or smoke a cigarette in the future or belief that will get drunk before age 16 years.⁶⁰

The strongest evaluation in this category, that by Flay *et al.*⁴⁴ of the AAYP intervention, reports the least promising results of school environment intervention. However, although not a focus of this review,

compared with the no-intervention comparison for boys, both the school environment plus curriculum arm and the curriculum-only arm provided significant benefits. The AAYP investigators themselves concluded that the environment component may have been critical to the overall impact of the combined environment/curriculum intervention based on the substantially larger effect size of the combined intervention than of the curriculum-only intervention, each compared with the no-intervention comparison, for example with regard to violent behaviour (47% compared with 35% reduction), drug use (34% compared with 32%) and recent sexual intercourse (65% compared with 44%).

Taken together these studies suggest the potential of school environment interventions to bring about benefits, particularly regarding measures concerned with violence and aggression. There was no evidence of any interventions causing harms. However, most of these studies were subject to important methodological limitations, such as not adjusting for clustering and high attrition. The strongest study in this category, that on the AAYP intervention, reported the most equivocal results, although the authors interpreted these as suggesting the potential benefits of school environment intervention. We conclude that the evidence base for this category of school environment intervention is promising but not definitive.

Two RCTs each reported on by a single report assessed interventions that combined changes to American middle schools' food and physical activity environments alongside actions that aim to empower students to contribute to achieving these changes. The studies were relatively well conducted, although the evaluation of the HYP intervention was subject to high and differential attrition. Both studies reported intervention benefits for some measures of student physical activity, but neither reported intervention benefits for any measures of student healthy eating. Although the evidence for the effectiveness of these interventions is therefore somewhat stronger than the evidence for interventions addressing sense of community and interpersonal relations, it is not possible to conclude whether or not such interventions to modify the school food/physical activity environments might have achieved the same effects even in the absence of empowered student participation. However, as reported above, mediation analysis in the HYP study did suggest that student empowerment may be important.

Three quasi-experimental evaluations reported in a total of five reports examined the effects of interventions to improve playgrounds in English elementary schools by either providing new marking and equipment or merely providing surface markings. These reported mixed findings regarding effects on students' physical activity. There were indications from one study that the benefits of improving playground physical structures may be greater for younger children and when recess was longer. The stronger of the two studies examining playground surface markings suggested that these were associated with significant effects on both VPA and MVPA immediately after intervention. Given the limitations of these three studies we must conclude that there is not currently a strong evidence base for the effects of these interventions.

Seven of the reports of intervention outcome evaluations examine subgroup effects but only one examined effects by SES. Furthermore, subgroup reporting was carried out inconsistently and three reports failed to report tests for interaction so these studies provide little information on the likely impact of school environment interventions on health inequalities. Only the AAYP intervention was delivered in an explicitly deprived area; this study supports the notion that school environment interventions addressing low-income communities are feasible.

None of the evaluations reports on cost-effectiveness. With the exception of the playground interventions reported by Ridgers *et al.*¹²⁴ and Stratton and Mullan,¹¹⁸ none of the studies reported on intervention costs. The descriptions of the interventions do not provide enough data to enable estimates of costs, particularly in terms of the costs for schools that arose as a result of their participation in the interventions.

Our focus is interventions that aim to modify the school environment without simultaneously addressing school health curricula. We have concluded that there is evidence for the potential of such action to promote

student health, but that the evidence is not currently definitive. We will assess the implications of these results for our review hypotheses in *Chapter 11*.

These results should not be taken as suggesting that there is no definitive evidence for the effectiveness of health-promoting school interventions or co-ordinated school health programmes, which, as well as addressing school ethos or environment, include curricular and community components. A Cochrane review of the effects of health-promoting school interventions comprising changes to school environment/ethos, curricula and parent/community engagement is proceeding in parallel to our own review and was due to report later in 2012. To develop a fuller picture of the effects of school environment interventions, the results of our own review should be read in conjunction with those of the Cochrane review.

Strengths and limitations

This section addresses limitations that are particular to the review of outcome evaluations; limitations that apply across our different reviews are considered in *Chapter 11*. Our review excluded interventions that included health education alongside actions to address the school environment to be able to assess the specific effects on health of modifying the school environment and in particular how schools are organised, managed and built and how they provide teaching, pastoral care and discipline. However, some of the intervention studies that fit our inclusion criteria nonetheless did not quite lend themselves to these ends. The studies of whole-school approaches to promoting physical activity and healthy eating reported by Dziewaltowski *et al.*⁶³ and Sallis *et al.*¹²² were included because they sought to enable student participation in decision-making (fitting our criterion concerning pastoral care) and did not include health education, but they also included actions altering the school food environment and opportunities for physical activity, which was not our focus. We could not establish whether or not any outcomes could be attributed to student participation, although mediation analysis in one paper suggested that this was likely.

Our criteria allowed us to make clear decisions about which study reports to include or exclude, although we sometimes needed to make difficult judgements. For example, we had to make judgements about whether or not interventions should be considered as school environment interventions. Some classroom management interventions were excluded. For example, an evaluation of the Good Behaviour Game was excluded because of its use of a highly structured classroom activity to improve students' behaviour that was delivered in certain specific sessions, which we therefore took to be a form of health education curriculum. In contrast, we included the CDP intervention because this addressed classroom management in a pervasive rather than a discrete, modularised manner, as per the Good Behaviour Game. Although the CDP intervention was not delivered in all of the classrooms in each school, we regarded it as nonetheless aiming to transform a school's overall environment, and the RCT evaluating it involved the allocation of schools not classrooms to intervention or control arms. We also excluded a study by Twemlow *et al.*¹²⁷ because the intervention included some martial arts training that includes self-regulation, role-play and anger management (which we regarded as a health education curriculum). In contrast, we included the study by Fonagy *et al.*,¹²⁰ which also involved self-defence training, because it was geared towards equipping students with the skills to deliver the intervention and did not include any reference to broader personal or social skills development.

Chapter 8 Research question 3: process evaluations

Research question

How feasible and acceptable are the school environment interventions examined in studies addressing RQ2? How does context affect this, examined using process evaluations linked to outcome evaluations reported under RQ2?

Methods

Inclusion and exclusion criteria

As reported in the last chapter, 16 reports were included that addressed RQ2 (outcome evaluations). We included process evaluations linked to outcome evaluations reviewed in the last chapter. To identify process evaluations associated with these, the full texts of these reports were retrieved and the following exclusion criteria were applied by one reviewer and checked by a second reviewer (there were no discrepancies to be resolved):

- exclude if study is not a process evaluation
- exclude if study does not report on an intervention subject to an outcome evaluation included in stage 2
- exclude if study is not written in English.

Quality assessment

All included reports were quality assessed using the following criteria:

- whether or not study has clear RQs/aims
- whether or not sampling and sample are described
- whether or not study examines planning (using qualitative data)
- whether or not study examines delivery (using quantitative or qualitative data)
- whether or not study examines coverage (using quantitative or qualitative data)
- whether or not study examines receipt (using quantitative or qualitative data)
- whether or not study examines acceptability (using quantitative or qualitative data)
- whether or not study examines context (using quantitative or qualitative data).

These criteria used for assessing methodological quality were adapted from those used in a previous review.⁹⁴ They allowed us to assess which studies were well reported, which examined the intervention process comprehensively and which enabled examination of the process from a range of perspectives. Reports were not excluded or graded based on these quality assessment ratings. Instead, this assessment was used qualitatively when weighing up evidence from each evaluation.

The criteria were piloted on a random sample of two reports by two reviewers (CB and HW) before being applied by one reviewer (HW) and checked by another (CB), with any differences being settled by discussion without recourse to a third reviewer.

Data extraction

Because only those process evaluation studies that were linked to included outcome evaluations (RQ2) were included, we had already extracted data on study RQs/hypotheses, study site and population, sampling, data collection methods, analysis methods and results. Informed by existing tools for data extraction of

process evidence^{126,128} we extracted data related to the following: part of the process examined (planning, delivery, receipt), aspect of the process examined (feasibility, fidelity, coverage/accessibility, acceptability) and aspect of the intervention context examined (e.g. measured need, policy, institutional and professional capacity, collaboration, 'product champions'). Data extraction tools were piloted on a random sample of two reports by two reviewers (CB and HW) before being applied by one reviewer (CB) and checked by another (HW), with any differences being settled by discussion.

Synthesis

A narrative synthesis was conducted for process evaluation studies. We included all studies in the narrative, but made clear where studies were subject to methodological limitations informed by our quality assessment. It was aimed that the narrative synthesis would develop overarching themes, but in practice it was largely restricted to narrative summaries of the findings of each study in context. This was because the studies were too heterogeneous in design and methods to develop meaningful or very detailed overarching narrative themes. Nevertheless, a narrative overview is provided in this chapter's discussion section.

Overview of included reports

Flow of literature

Only those process evaluations that were linked to outcome evaluations included in RQ2 were considered for inclusion. Therefore, the 16 included outcome evaluation study reports were screened for accompanying process evaluations (all of the reports were written in English). Of these 16, five reports included process evaluations. From checking the references of the five included reports, we identified one further linked process evaluation that was included in the in-depth synthesis. Fagen and Flay¹²⁹ reported on a process evaluation of the sustainability of the AAYP intervention but this focused only on the curriculum component and so is not considered further here. Thus, six reports (one linked)^{55,58-60,84,85} of four studies were included in the process evaluation in-depth synthesis (*Figure 9*).

Quality assessment

Study reports varied in whether or not they set clear RQs. Clear questions were provided by Bonell *et al.*^{59,60} and Solomon *et al.*⁵⁸ Nearly all described their sampling methods and samples; Solomon *et al.*⁵⁸ did not describe the sample. Studies varied in the extent to which they sought the perspectives of a range of stakeholders on the interventions. Battistich *et al.*⁵⁵ and Solomon *et al.*⁵⁸ relied solely on research observations of delivery. Bonell *et al.*^{59,60} collected data from external providers, school staff and students. Dzewaltowski *et al.*⁶³ collected data from students, site co-ordinators and teachers. Flannery *et al.*¹¹⁹ collected

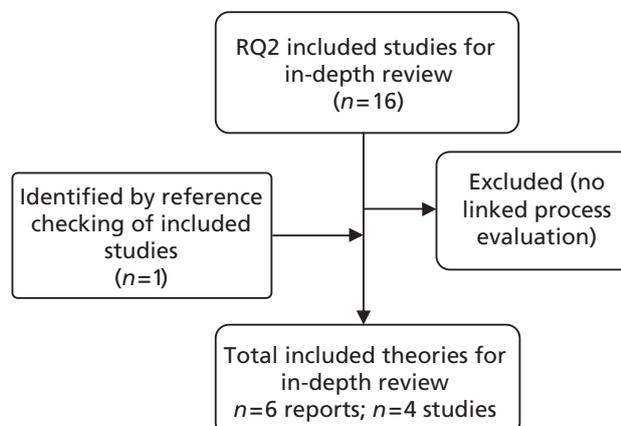


FIGURE 9 Flow of literature: process evaluation synthesis (stage 2: in-depth synthesis).

data solely from teachers. Only Bonell *et al.*^{59,60} drew on qualitative and quantitative data, the others drawing only on quantitative data.

Only Bonell *et al.*⁵⁹ examined the planning of the intervention prior to delivery, drawing on interviews with providers. Reports from all four interventions described the fidelity of intervention delivery using observations or data from teachers, other intervention staff or students. Only Battistich *et al.*⁵⁵ and Solomon *et al.*⁵⁸ reporting on the CDP intervention, examined the fidelity of delivery by comparing intervention with control schools. Bonell *et al.*^{59,60} and Dzewaltowski *et al.*⁶³ drew on self-report data from providers to quantify delivery. The HSE evaluation^{59,60} also drew on qualitative observational data on delivery and interview data on the feasibility of delivery, while Flannery *et al.*¹¹⁹ examined quantitative data from teachers on feasibility. Bonell *et al.*^{59,60} and Dzewaltowski *et al.*⁶³ examined coverage by determining recognition of the intervention among students. Bonell *et al.*^{59,60} evaluated acceptability qualitatively through interviews and focus groups. Flannery *et al.*¹¹⁹ evaluated acceptability quantitatively through surveys only of the training component for teachers. Dzewaltowski *et al.*⁶³ examined the extent to which training provided site co-ordinators with the self-efficacy to undertake their work. Only Bonell *et al.*^{59,60} drew on qualitative data and examined how context might influence intervention delivery or uptake.

Study characteristics

Five reports examined three interventions that encouraged staff and students to develop school climates characterised by a stronger sense of community and better relationships.^{55,58-60,119} One study evaluated an intervention that enabled staff and students to advocate for school environments promoting healthier eating and physical activities.⁶³

Results

Narrative summary of findings from each process evaluation

Battistich *et al.*⁵⁵ report that, across all years of the CDP intervention, teaching practices across the five areas addressed by the programme were significantly distinctive from those in comparison schools, suggesting that the intervention was feasible to deliver with good fidelity. Solomon *et al.*⁵⁸ report that according to observations of teachers' there were significant differences between intervention and control schools in around half of the teaching practices and classroom activities intended to be brought about by the project, for example teachers' use of group praise and students participating in rule development. According to student reports, there were significant differences between intervention and control arms in most indicators of teaching practices and classroom activities prescribed by the project. This study examined intervention fidelity in a rigorous manner so that conclusions in this area are likely to be sound, but did not examine other aspects of or perspectives on process, or assess context.

Bonell *et al.*⁵⁹ report that the intervention was delivered as intended with all components implemented, although it should be noted that this study did not examine fidelity of delivery through observation sessions. Qualitative data suggest that the external facilitator enabled schools to convene an action team involving staff/students. Inputs were feasible and acceptable and enabled similar actions in both schools. Locally determined actions (e.g. peer mediators) were generally more feasible and acceptable than preset actions (e.g. modified pastoral care). This study alone used qualitative data to examine the effect of contextual factors on implementation. This suggested that implementation was facilitated when it built on aspects of schools' baseline ethos (e.g. a focus on engaging all students, formalised student participation in decisions) and when senior staff led actions, acting as 'product champions'. Student awareness of the intervention was high. Quantitative data on students' attitudes and behaviours suggested that the intervention aims corresponded with local needs in each intervention school. Bonell *et al.*⁶⁰ report that some activities such as rewriting school rules involved broad participation, which was assessed through qualitative methods such as interviews.

Flannery *et al.*¹¹⁹ report that teachers being trained to deliver the PeaceBuilders intervention found the philosophy behind the intervention easy to understand. They regarded training as clear, effective and easy to

follow, that the school administration supported the intervention and that it would be easy to implement and would be effective in the classroom. Surveys with teachers also suggested that the intervention was delivered regularly; approximately half rated implementation as extensive and half as moderate, with around half using half or more of the intervention materials. Other aspects of the process, other perspectives, such as those of students, and context were not examined.

Dzewaltowski *et al.*⁶³ undertook a thorough quantitative assessment of process and reported that training for the HYP intervention occurred as planned, with site co-ordinator attendance very high in both years and self-efficacy arising from training being high. Site co-ordinators formed 'change teams' in each school that met regularly. Site co-ordinators reported an average of 26.5 implemented programme, policy or practice changes. Teachers implemented around two-thirds of planned lessons. Student surveys suggested that around one-third of students had heard about the intervention or its activities, around half of whom had participated on 'change teams'. Context was not examined in this study.

Discussion

Narrative overview of findings

Of the 16 included outcome evaluations, five reports included process evaluations and one further linked process evaluation paper was found by reference checking. These employed a range of research methods, most frequently drawing on quantitative data collected from students and/or teachers. These reports addressed some aspects of our third RQ more than others. Although most examined feasibility or fidelity in some way, fewer examined acceptability and only one study used a mix of quantitative and qualitative methods to examine local context and how this influenced intervention processes. Process evaluations reported largely positive results regarding intervention feasibility, fidelity, reach and acceptability, although differences in methods prevent any comparison of the delivery and uptake of each intervention. The single study that examined context suggested that this was important, facilitating implementation when this built on schools' existing ethos and when senior staff championed the intervention.

Strengths and limitations

We limited our in-depth review of process evaluations to those linked to outcome evaluation studies because in consultation with our stakeholders we deemed it most useful to synthesise evidence about the feasibility and acceptability of interventions about which we have evidence of their effects. This pragmatically limited the scope of our review by preventing us from synthesising evidence, for example, on potentially innovative interventions that have been feasibility tested but not yet been subject to outcome evaluation.

Most of our outcome evaluations were accompanied by process evaluations but these involved a diversity of methods making it impossible to compare the feasibility, coverage, acceptability or context of the interventions. The small number of heterogeneous studies made it impossible to draw conclusions about how context, processes and outcomes might inter-relate.

Chapter 9 Research question 4: multilevel studies

Research question

What are the effects on health and health inequalities among school students aged 4–18 years of school-level measures of school organisation and management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers, and/or the physical environment (measured using 'objective' data), examined using multilevel quantitative designs?

Methods

Inclusion and exclusion criteria

There were 285 multilevel and ecological references coded in the evidence map. We included multilevel studies examining the effects on student health or well-being outcomes (age 4–18 years) of objective school-level measures of how schools are led and managed, how they teach, support or discipline students, or schools' physical environment. Two reviewers (CB and HW) double-sifted all 285 references, independently, using the exclusion criteria shown in *Table 5*.

We screened all references identified as potentially relevant to RQ4 in the evidence map based on title and abstract only and full paper where necessary, not just those coded as being about relevant school-level exposures, because we could not assume that coding was accurate for the evidence map. Screening was not hierarchical or mutually exclusive, so some reports were excluded based on multiple criteria.

We included only multilevel studies and not ecological studies. Ecological studies measure the prevalence of outcomes at the school level and relate these to school-level characteristics. For example, they might conclude that smaller schools or schools with a high staff–student ratio have lower rates of smoking. Multilevel school studies measure outcomes at the individual level and explain these in terms of school-level characteristics as well as individual-level student characteristics. For example, they might conclude that student smoking is affected not only by the quality of the individual student's own relationships with teachers but also by the overall quality of relationships between students and teachers in a school. Ecological studies are vulnerable to unmeasured confounding, whereby differences in outcomes between schools reflect differences in composition rather than differences in school environments, and are also unable to disentangle the effects of school-level factors such as student–staff relations that can also be represented at the individual level (cross-level bias).⁴⁸

We decided to focus only on those studies in which measures of the school environment drew on 'objective' data', that is, data were not derived from aggregates of self-reports from the same individuals (usually students) who also provided data on health outcomes. This was a pragmatic decision to limit the number of included reports as well as being informed by our view that, when data on exposures and outcomes were derived from the same individuals, any associations found might merely be the result of unmeasured characteristics of individuals unrelated to the effects of the school environment.

Quality assessment

All included reports were quality assessed using the following criteria, which were informed by the review by Aveyard *et al.*:⁴⁸

- whether or not appropriate covariates (key potential confounders) were adjusted for
- whether or not inappropriate covariates (potential mediators) were adjusted for.

TABLE 5 Exclusion criteria for multilevel studies of school effects on health (in-depth synthesis)

Exclusion criterion	Guidance
Exclude 1: not school-level measures	Exclude if the study does not report on the health effects of school-level measures of school type, school leadership and management, teaching, support and discipline, student attitudes to school or relations with teachers, and/or physical environment
Exclude 2: no objective measures	Exclude if objective measures were not used
Exclude 3: student aggregate data only	Exclude if student aggregate data (i.e. survey data) were used
Exclude 4: not a multilevel model analysis	Exclude if study does not involve multilevel analysis of school-level and student-level determinants of health outcomes
Exclude 5: no relevant health outcomes	Exclude if there are no relevant health outcomes
Exclude 6: other reason	Exclude for any other reason (e.g. teacher rather than student health; topic; exposure not concerned with school management)
Exclude 7: non-English	Report is written in a language other than English
Include	Study meets the criteria for in-depth synthesis

By confounders we are referring to factors such as students' SES or sex, which could vary between schools and could themselves affect health outcomes, but which are unlikely to be influenced by schools during the time period of the research. By mediators we are referring to factors such as student engagement and peer smoking, which might affect health outcomes but which themselves might be open to school-level influences. Reports of multilevel studies were not excluded from in-depth review based on judgement of their quality. We included all of the reports meeting the inclusion criteria for stage 2 when describing study context and characteristics as well as when summarising intracluster correlation coefficients for the various health outcomes reported by multilevel studies. However, we did restrict our narrative synthesis of the effects of school-level determinants on these health outcomes to studies that we judged had taken an appropriate approach to adjusting for confounding. Adjusting for confounding is important to minimise bias in estimates of the association between school-level factors and student health outcomes. Differences in the rates of health outcomes between schools might sometimes be merely attributable to differences in the characteristics of students entering the school, their families or the neighbourhoods they live in. Therefore, it is important that studies adjust or otherwise control for these potential confounders. However, as some authors have pointed out,^{48,49,130} it is also important that studies do not overadjust for covariates that might lie on the causal pathway underlying this association. Covariates in this category would include any factors that might be subject to school-level effects, such as attitude to school, peer behaviours and academic attainment.

First, we judged which covariates were inappropriate to adjust for. It was not possible to develop a comprehensive list a priori because of the large range of covariates that we would encounter. However, we were informed by the following lists suggested previously by Aveyard *et al.*⁴⁸ and used these to aid our own case-by-case judgements about which covariates are inappropriate to adjust for because they might plausibly be influenced by the school.

Factors categorised into those influenced by school and those not influenced by school – appropriate to adjust for:

- family characteristics
- family structure
- parental smoking
- parental attitudes
- sibling smoking
- family environment
- attachment to family.

Pupil characteristics not influenced by school – appropriate to adjust for:

- sex
- ethnicity
- SES
- personal income.

Pupil characteristics potentially influenced by school – inappropriate to adjust for:

- age
- school engagement
- other risk behaviour
- stress
- depression
- self-esteem
- attitudes to smoking
- attitudes towards health generally
- peer smoking
- peer attitudes and norms.

Second, we examined whether or not studies did adjust for the key potential confounders. We set two thresholds of varying strictness. This was a pragmatic decision to enable us to focus on the least biased studies while ensuring that we excluded only a few studies. Our broader, more inclusive definition required that, as well as not adjusting for any inappropriate covariates, studies adjusted for sex plus some measure of individual and/or local area social disadvantage. We treated individual or family SES or ethnicity as well as family structure and local measures of deprivation or local health indices as being crude indicators of individual or area social disadvantage.

Our stricter definition was that, as well as not adjusting for any inappropriate covariates as defined above, studies adjusted for some measure of sex plus some measure of student ethnicity, family SES or structure plus some measure of area deprivation or local health indices. We set this stricter threshold because we judged that individual and area social disadvantage would operate as independent sources of confounding.

One reviewer (WP) applied these quality assessment criteria and another reviewer (CB) checked these assessments, with any differences being settled by discussion.

Data extraction

For the reports of multilevel studies we extracted data on study RQs/hypotheses, study site and population (i.e. overall baseline characteristics), data set, sampling, types of schools, data collection methods, analysis methods, results and authors' conclusions. We also extracted data on the objective school-level measures included in the multilevel models, the types of levels included (i.e. student, class, teacher, school, city, region) and the types of covariates (and whether appropriate or inappropriate). Data extraction tools were piloted on a random sample of two reports by two reviewers (CB and HW). One reviewer (HW) extracted data on context and methods of data collection and another reviewer (WP) extracted data on methods of analysis and results, both checked by a third reviewer (CB), with any differences being settled by discussion.

Synthesis

We assessed the potential for statistical meta-analysis of multilevel studies of school effects on health by noting which studies focused on similar combinations of school-level interventions/exposures and outcomes. We set a threshold of a minimum of three such studies being required to consider meta-analysis as meta-analysis of only two studies is unlikely to produce stable effect estimates.¹¹⁶ Among the included multilevel modelling studies we found possible homogeneity for Evans-Whipp *et al.*¹³¹ and Piontek *et al.*¹³² in their examination of tobacco policy comprehensiveness (although with quite different measures) and

smoking in the last 30 days; and Aveyard *et al.*⁶⁴ and Markham *et al.*⁶⁵ in their examination of value-added scores as well as truancy in relation to weekly smoking. We found no homogeneous triplets of studies. We therefore subjected our studies of school-level health effects to narrative synthesis only.

We restricted our narrative synthesis to those studies that we judged as taking an approach to adjusting for covariates. First, we developed narrative summaries of the findings reported by each study in the context of study quality, setting and population. This is reported in the results section of this chapter. Second, we developed a narrative overview of findings. We noted where studies of similar influences reported convergent or divergent findings and considered the reasons for this in terms of design and context, rather than merely assessing the overall frequency of differing findings. This is reported in the discussion section of this chapter. Third, we considered the implications of the results of multilevel studies of health effects alongside other forms of evidence for this review's primary and secondary hypotheses. This is reported in *Chapter 11*.

Overview of included reports

Flow of literature

Of the 285 references identified from the evidence map, 236 were excluded based on the criteria in *Table 5*. In addition, five reports were duplicates and the full text was not available for six of the reports. A further four reports were identified from citation chasing of reports included in the evidence map. Thus, 42 reports (from 34 data sets) were included in the multilevel studies in-depth synthesis (*Figure 10*).

Quality assessment

Adjustments for inappropriate covariates

The covariates shown in *Table 34* in *Appendix 6* were considered plausibly to lie on a causal mechanism between school-level exposures and student health outcomes and were therefore judged as inappropriately adjusted for.

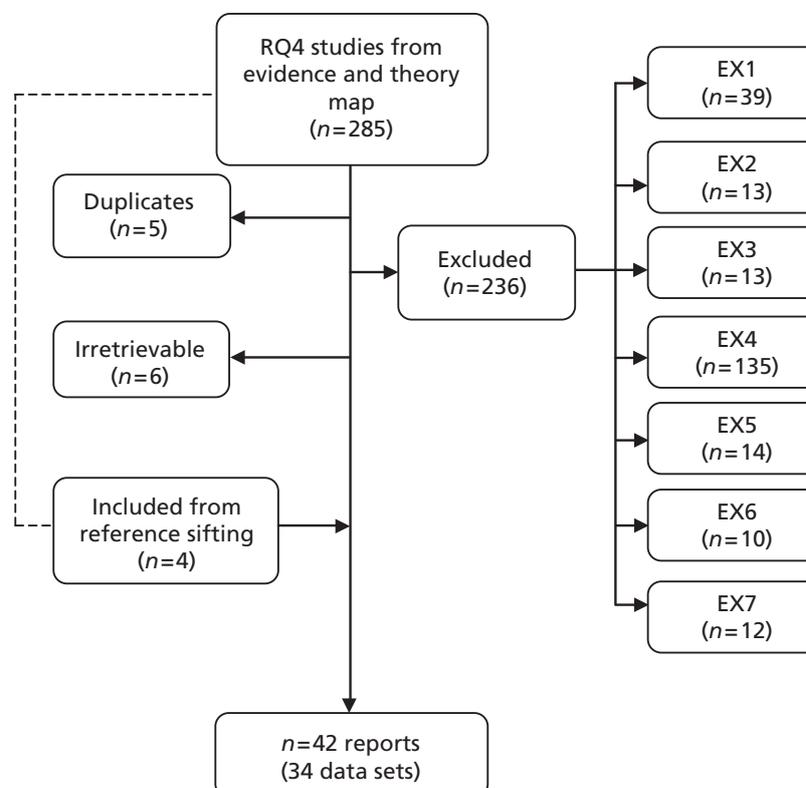


FIGURE 10 Flow of literature: multilevel studies (stage 2: in-depth synthesis).

Adjustments for appropriate covariates

The covariates shown in *Table 33* in *Appendix 6* were appropriately adjusted for.

Taking our broader definition of adjustment for the key appropriate confounders, the following studies adjusted for the key appropriate confounders and did not adjust for any inappropriate covariates: Aveyard *et al.*,⁶⁴ Bisset *et al.*,⁶⁶ Cradock *et al.*,¹³³ Evans-Whipp *et al.*,¹³¹ Henry *et al.*,⁶⁷ (some models), Kumar *et al.*,⁶⁸ Markham *et al.*,⁶⁵ Monshouwer *et al.*,⁶⁹ Piontek *et al.*¹³² and Tobler *et al.*⁵³

Taking our stricter definition of adjustment for the key appropriate confounders, only a handful of studies would be regarded as adjusting for key appropriate confounders and not adjusting for any inappropriate covariates: Aveyard *et al.*,⁶⁴ Bisset *et al.*,⁶⁶ Evans-Whipp *et al.*,¹³¹ Henry *et al.*,⁶⁷ Markham *et al.*⁶⁵ and Tobler *et al.*⁵³

Study characteristics

Overall, 17 reports focused on US samples,^{53,67,68,70–73,76,77,81,82,85,86,88,89,133,134} six were from Canada^{79,80,83,84,90,135} and six were from the UK,^{48,65,66,130,136,137} three were from Norway^{74,75,87} two each were from Israel^{78,138} and the Netherlands^{69,139} and one each was from Australia,⁹⁷ Belgium,¹⁴⁰ Germany,¹³² Spain,¹⁴¹ and Thailand,¹⁴² with one additional study¹³¹ reporting on data from the USA and Australia.

Two reports did not indicate what kind of schools were researched^{70,71} and five reports indicated only from which grades research participants were drawn (without reporting schools' full grade structure).^{72,75}

Of the other studies:

- 18 focused on secondary/high schools^{65,66,69,77,79,81–84,87,97,131,132,136,138,139–141}
- three focused on middle and secondary/high schools^{48,76,134}
- four focused on primary/junior/elementary schools^{86,90,135,162}
- one focused on primary/junior/elementary and middle schools⁸⁹
- three focused on primary/junior/elementary and secondary schools^{80,130,137}
- two focused on middle schools^{85,133}
- one focused on high/junior high.⁶⁷

Aveyard *et al.*⁶⁴ and Bisset *et al.*⁶⁶ both drew on data from the West Midlands Young People's Lifestyle Survey, but examined different outcomes. Two reports by Wium and Wold^{74,75} both drew on Control of Adolescent Smoking (CAS) data, but examined different smoking outcomes. Two reports by West *et al.*^{130,137} both drew on the West of Scotland 11–16 survey, but examined different outcomes and the latter involved more follow-ups. Gastic⁷⁶ and Haley⁷⁷ both drew on the National Longitudinal Study of Adolescent Health (NELSAH or 'Add Health') data, but used differing waves of follow-up and examined different outcomes. Attar-Schwartz and Khoury-Kassabri¹³⁸ and Khoury-Kassabri *et al.*⁷⁸ both drew on the same Israeli data set, but examined different outcomes. Kairouz and Adlaf⁷⁹ and Lothian⁸⁰ both drew on the Ontario Student Drug Use Survey (OSDUS) data set and also examined different outcomes. Kim⁸¹ and Way⁸² both drew on National Education Longitudinal Study (NELS) data, but the latter used data from more follow-ups. Ishibashi⁸³ and Murnaghan *et al.*⁸⁴ both drew on data from the School Health Action Planning and Evaluation System (SHAPES) study, but used different waves to examine different outcomes.

Other methodological issues

Although not the focus of our quality assessment, a range of methodological issues was identified by our review of studies.

Poor reporting

Most of the studies failed to report *all* of the information sought in this review. For example, seven reports^{64–66,81,133,134,137} failed to report basic descriptive analyses of mean prevalence or overall average of the outcomes modelled. Eighteen studies^{53,64–66,68,71,84–88,131,134,139,142,143} failed to report the intraclass correlation of outcomes for a null model (i.e. a model containing no covariates).

Twenty-five studies^{53,65,68,69,71,72,74,75,77,81,83–85,87,89,90,97,130–135,139,142} also did not report the proportion of between-school variance explained by school-level variables in the models.

Three of the doctoral dissertations/theses^{81,85,86} and one research study⁸⁷ were particularly poorly reported. In these cases it was difficult to determine which variables were included in particular multilevel models.

Three studies^{72,80,135} suffered from problems of reporting caused by the use of determinants whose scale of measurement was several orders of magnitude larger than that of the outcome. In these cases, even when significant associations were estimated, the reported coefficients (and any corresponding ORs) were of no practical use because they were correspondingly very small.

Response rate

Many of the studies reported non-trivial proportions of missing data in their samples. This does not necessarily imply that the resulting analyses will be biased. Randomly missing data (missing completely at random, termed MCAR) does not cause biases in analyses, and biases from systematically missing data can often be accounted for by including covariates in the model that are associated with non-response propensity, reweighting the available sample or imputing missing items (in the case of data missing conditionally at random, termed MAR). It is therefore good practice to analyse missing data patterns and to try to adjust for any identified biases. Haley⁷⁷ weighted his data to match the original population, but identified a bias due to missing items. He ignored this bias in the modelling, using only cases with complete data. Haug *et al.*⁸⁷ used multiple imputation to fill in some of the missing data at the school level. Johnson and Hoffmann⁷¹ also used this approach to fill in all missing items. Khoury-Kassabri *et al.*,⁷⁸ Kim,⁸⁸ Kumar *et al.*⁶⁸ and Monshouwer *et al.*⁶⁹ used weights to make their samples more representative of the populations they were modelling. West *et al.*¹³⁷ reported models based on unweighted data after weighting for attrition resulted in unchanged estimates. Where imputation had been used, little information was provided as to the specification of imputation models. Imputation and reweighting methods are only as effective as the quality of the data and methods used to implement them. Overall, it is not clear to what degree missing data in the studies led to model estimates suffering from bias. In any case, the majority of studies reported complete-case analyses, effectively ignoring potential bias due to missing data.

Longitudinal design

Nine studies^{65,72,76,81,82,86,130,137} involved longitudinal designs with repeated measurements of young people being made over time. In most cases sample attrition was relatively low, but in Tobler *et al.*⁵³ only 61% of the initial cases were followed up, and in the reporting of Kim⁸¹ and Way⁸² it is not clear how much attrition of the samples occurred. In West *et al.*,¹³⁷ final follow-up (at wave 4) is only 45%, but follow-up rates were much better in earlier waves.

Only in the two reports by West *et al.*^{130,137} did the modelling involve baseline measures from before the young people entered the school forming the focus of the study. This means that only these two reports adjusted for true pre-exposure characteristics of the young people (e.g. baseline health behaviours in West *et al.*,¹³⁰ maths ability in West *et al.*¹³⁷).

None of the longitudinal studies seemed to involve the use of three-level designs, in which measures at each wave are clustered in individual young people who are then clustered in schools. This implies that there is a possibility of heterogeneity at the individual level being conflated with the estimated coefficients of model covariates. Likewise, none of the models, except for that of Johnson and Hoffmann,⁷¹ can be said to have explicitly examined longitudinal health behaviour transitions. Instead, they actually modelled mean health behaviour prevalence over time.

Statistical issues

Both the Rountree and Clayton study⁷⁸ and Way's study⁸² suffered from serious model misspecifications, whereby continuous regression approaches were used with ordinal outcome variables. Rountree and Clayton⁷³ also entered covariates into their model as continuous when they were ordinal, another flaw in

their model specifications. A slight concern was also noted in the model specification of Piontek *et al.*¹³² Despite having a categorical variable for school type entered into the model, only one coefficient was reported, whereas this should have resulted in three dummy coefficients for the four category variable.

School-level measures

Nine reports^{68,69,71,76,77,82,130,132,139} examine school type in terms of public/private or denominational. Transition structure is examined by Henry *et al.*⁶⁷ and West *et al.*¹³⁷ School size is examined by 17 study reports^{67,68,71–73,76–80,82,83,86,130,135,137,138}. The following reports examine class size: Attar-Schwartz and Khoury-Kassabri,¹³⁸ Khoury-Kassabri *et al.*,⁷⁸ Gastic⁷⁶ and Kim.⁸⁸

The only reports to examine schools' physical environment are Cradock *et al.*¹³³ (campus area per student, playground area per student and school building area per student), Haug *et al.*⁸⁷ (availability of facilities for physical activity), Kumar *et al.*⁶⁸ (attractive physical environment, neglected physical environment and the total number of unobservable and unsupervised places in and around school) and Malikaew *et al.*¹⁴²

Eight reports examine aspects of school staffing: Henry *et al.*,⁶⁷ Johnson and Hoffmann,⁷¹ Way⁸² and Bradshaw *et al.*⁸⁹ (student–teacher ratio); Akiba⁷⁰ (certification of teachers); Grunseit *et al.*⁹⁷ (teachers' experience); Maes and Lievens¹⁴⁰ (whether or not school administrator is female, high teacher workload, sex ratio of teachers); and Wiium and Wold⁷⁵ (number of years the teachers' smoking policy has been in place and whether or not teachers were involved in its development).

Akiba⁷⁰ examines the effects of academic tracking. Aveyard *et al.*,⁶⁴ Bisset *et al.*,⁶⁶ Tobler *et al.*⁵³ and Markham *et al.*⁶⁵ examine the effects of educational attainment, gross and value added. Haley⁷⁷ reports on rates of dropouts.

The effects of smoking and alcohol rules are examined by nine reports.^{69,74,75,84,131,132,134,136,141} Measures of other aspects of school discipline are examined by Comeau,⁸⁵ Gastic,⁷⁶ Haley,⁷⁷ Kim,⁸⁸ Leatherdale *et al.*⁹⁰ and Maes and Lievens.¹⁴⁰ Bradshaw *et al.*⁸⁹ report on the effects of rates of suspensions. Two reports^{76,85} examine the effects of observed/reported school misbehaviour. Two reports^{64,66} examine the effects of truancy rates. Three reports examine some measure of ethos.^{85,130,137} Two reports^{81,137} examine the effects of parent-related policies. Two reports^{87,137} look at other aspects of school policies.

The way in which the school-level measures examined in studies included in our narrative synthesis map against the domains specified in our review hypotheses is summarised in *Chapter 11* (see *Table 9*).

Health outcomes

Smoking-related outcomes were examined in 17 reports.^{53,64,65,68,71,74,75,80,84,88,130–132,134,136,140,141} Alcohol outcomes were examined in 12 reports.^{53,66–69,73–75,77,79,88,140} Illicit drug or solvent use was examined in six reports.^{53,66,68,72,74,88} Overall substance use was examined by Comeau.⁸⁵

Fear of violence was reported by Akiba,⁷⁰ Bradshaw *et al.*,⁸⁹ Gastic⁷⁶ and Gladden.⁸⁶ Verbal or physical victimisation outcomes were reported by Attar-Schwartz and Khoury-Kassabri¹³⁸ and Bradshaw *et al.*⁸⁹ Verbal or physical aggression was reported by Grunseit *et al.*,⁹⁷ Khoury-Kassabri *et al.*⁷⁸ and Kim.⁸⁸ Fighting-related outcomes were reported by Gladden,⁸⁶ Haley,⁷⁷ Tobler *et al.*⁵³ and Way.⁸²

Physical activity-related outcomes were examined in five reports.^{83,87,90,133,139} Diet-related outcomes were reported by Ma,¹³⁵ van der Horst *et al.*¹³⁹ and Wiium and Wold.⁷⁴ A weight-related outcome was reported by van der Horst *et al.*¹³⁹

Dental-related outcomes were reported by Maes and Lievens¹⁴⁰ and Malikaew *et al.*¹⁴² Physical health was reported by Ma.¹³⁵ Emotional and mental health-related outcomes were reported by Ma¹³⁵ and West *et al.*¹³⁷

Results

Narrative summaries of intraclass correlation coefficients of health outcomes

It is generally accepted that the intraclass correlation coefficient (ICC) of an outcome under the null model (a model containing an intercept term and no covariates) provides a good descriptive measure of the degree of clustering in data. More formally, the ICC is the proportion of total outcome variance that is 'between groups', that is, in this review the proportion of total outcome variance that is associated with the school level.

In reviewing study findings on ICCs we draw on all reports regardless of the approach taken to adjustment for covariates because the ICC data were drawn from unadjusted models. Reported ICCs were generally quite low, with most falling in the range of 0–10%, indicating that most of the variation in outcome measures was due to variability at the student level as opposed to the school level. Notable exceptions to this were verbal victimisation (ICC=21%) in the study by Attar-Schwartz and Khoury-Kassabri¹³⁸ and healthy food eating (ICC=17%) and exercise (ICC=29%) in the study by Ma.¹³⁵

Some studies report ICCs for different measures and so allow comparison of the apparent clustering of different outcomes. Attar-Schwartz and Khoury-Kassabri¹³⁸ and Khoury-Kassabri *et al.*⁷⁸ reported that verbal victimisation was more clustered than indirect victimisation; Collins *et al.*⁷² reported that marijuana use was more clustered than other drug use; Haley⁷⁷ reported that drinking among same-sex-attracted students was more clustered than fighting; Kim⁸¹ reported that clustering for alcohol use was higher than that for smoking; and Ma¹³⁵ reported that clustering was greatest for exercise but lower for healthy eating and even lower for mental health. However, these findings from single studies cannot provide a clear sense of underlying patterns of clustering for these outcomes. Because of the wide variation in outcome measures, study contexts and data collection methods, direct comparisons of ICCs across studies, even where outcomes were substantively similar, would not be valid.

Narrative summary of the results of well-adjusted studies on school-level effects

The model estimates for determinants reported in studies that were well adjusted are outlined in the following sections. We will start with the reports that met the narrower definition of adjusting for all key confounders before moving on to those that met the broader definition.

Studies meeting the narrower definition of well adjusted

Six studies^{53,64–67,131} met the narrower definition of well adjusted.

In the Aveyard *et al.*⁶⁴ cross-sectional study of students aged 11–16 years in secondary schools in the West Midlands, UK, there was no significant association of school achievement [five good General Certificate of Secondary Education (GCSE) passes based on a 5-year average, 1994–8] with regular smoking. Similarly, there was no association of school truancy (based on a 5-year average of half-days lost through pupil absence, 1994–8) with pupils' regular smoking. However, there was an association of regular smoking with whether a school was classified as 'authoritative' or 'laissez faire'. This categorisation was created by regressing school achievement and school truancy on five indicators of pupils' social profile, with the resulting school residuals then being used in a principal components analysis to create a single component measure of 'value added' by the school. Schools with a component measure one or more standard deviations (SDs) above average were named 'authoritative'; schools with a component measure one or more SDs below average were named 'laissez faire'. Authoritative schools had significantly lower (OR 0.83; 95% CI 0.73 to 0.95) and laissez-faire schools had significantly higher (OR 1.13; 95% CI 1.04 to 1.24) than average rates of regular smoking.

Bisset *et al.*⁶⁶ used the same cross-sectional data set as Aveyard *et al.*⁶⁴ No significant association was found between school achievement (the same measure as used by Aveyard *et al.*⁶⁴) and pupils drinking alcohol at least once a month. Higher levels of academic achievement of schools were associated with lower rates of heavy drinking (OR 0.93; 95% CI 0.89 to 0.97). There was no significant association of this achievement

measure with regular illicit drug use. There was no association of school truancy (the same measure as used by Aveyard *et al.*⁶⁴) with any of the following outcomes: pupils drinking alcohol at least once a month, heavy drinking or regular illicit drug use. As with Aveyard *et al.*,⁶⁴ a measure of 'value-added' school 'culture' was created. This was calculated by regressing school achievement and school truancy rates on five indicators of pupils' social profile with school residuals then being used in a principal components analysis to create a single component measure of value added, which accounted for two-thirds of the variance in the original truancy and achievement measures. This measure was then entered into the model as a continuous variable. This measure of value-added school culture was weakly but significantly associated with lower rates of drinking alcohol at least once a month (OR 0.87; 95% CI 0.78 to 0.95), lower rates of heavy drinking (OR 0.91; 95% CI 0.85 to 0.96) and lower rates of regular illicit drug use (OR 0.90; 95% CI 0.82 to 0.98). There was also a significant interaction of value-added school culture with school grade for rates of heavy drinking [χ^2 (2 degrees of freedom)=10.41; $p<0.01$], with lower grades being associated with lower rates of heavy drinking for the same level of value-added school culture (grade 7, OR 0.57; grade 9, OR 0.65; grade 11, OR 0.96).

Evans-Whipp *et al.*¹³¹ conducted a cross-sectional study of data from students aged 11–16 years in secondary schools in Washington state, USA, and Victoria, Australia. The study found no association of schools having a comprehensive smoking ban, harsh penalties for smoking, remedial penalties for smoking, smoking abstinence policies or smoking harm minimisation policies with the health outcomes: pupils smoking in the past 30 days, pupils smoking daily or perceptions of many students smoking on school grounds. One other determinant, strict enforcement of the smoking policy, did not have a significant effect on pupils smoking in the past 30 days or pupils smoking daily, but it was associated with reduced perceptions of many students smoking on school grounds (OR 0.45; 95% CI 0.25 to 0.82; $p<0.009$).

In a well-conducted, cross-sectional study of junior and middle high schools in rural USA that uses a novel and innovative analytical method called multilevel latent covariate modelling (MLCM), Henry *et al.*⁶⁷ reported that the following school-level factors were not associated with alcohol use among students aged 13–14 years: whether eighth graders are located within the same school as high school students or are in separate schools, school size and pupil–teacher ratio.

Markham *et al.*⁶⁵ reported from a longitudinal study of secondary schools in the West Midlands, UK (students aged 13–14 years at baseline) that a measure of 'value-added' education, similarly defined to that used in Aveyard *et al.*⁶⁴ and Bisset *et al.*,⁶⁶ was associated with a reduction of borderline significance in smoking at least one cigarette per week at first follow-up at age 14–15 years (OR 0.85; 95% CI 0.73 to 0.99 per SD increase in value added), and this became more significant at second follow-up at age 15–16 years (OR 0.80; 95% CI 0.71 to 0.91 per SD increase in value added). There was no significant interaction of value added with whether or not pupils were regular smokers at baseline.

Tobler *et al.*⁵³ report on a longitudinal study of middle school students in grade 6 at baseline (age 11–12 years) in the USA. Their main analytical model contained many appropriate covariates but it also contained an inappropriate covariate ('student academic problems') and so the base model was preferred. This study used a similar measure of 'value added' to that in Aveyard *et al.*,⁶⁴ Bisset *et al.*⁶⁶ and Markham *et al.*,⁶⁵ but calculated using 3-year averages. For follow-up in grade eight, when compared with schools in the range of –1 to +1 SD, value-added schools (≥ 1 SD above the mean) were associated with a lower incidence of drinking alcohol in the past 30 days (OR 0.60; 95% CI 0.42 to 0.88), drinking five or more alcoholic drinks on one occasion in the last 2 weeks (OR 0.44; 95% CI 0.23 to 0.84), smoking a cigarette in the past 30 days (OR 0.48; 95% CI 0.26 to 0.86) and using marijuana in the past 30 days (OR 0.29; 95% CI 0.15 to 0.57), as well as being involved in a group fight in the last month (OR 0.69; 95% CI 0.50 to 0.96). Value-attenuated schools (1 SD below the mean) did not have significantly different rates from those of other schools for any of the above outcomes.

Studies meeting the less narrow definition of well adjusted

Studies meeting the less narrow definition of well-adjusted and not adjusting for any inappropriate covariates were those by Cradock *et al.*,¹³³ Kumar *et al.*,⁶⁸ Monshouwer *et al.*⁶⁹ and Piontek *et al.*¹³²

In the Cradock *et al.*¹³³ cross-sectional study of middle schools in Boston, MA, USA, the school campus area per student (coefficient=0.2244; $p<0.001$), playground area per student (coefficient=0.347; $p<0.05$) and school building area per student (coefficient=2.1302; $p<0.05$) were all significantly associated with higher levels of physical activity during school hours, as measured using accelerometers among students aged 12–14 years.

In the Kumar *et al.*⁶⁸ cross-sectional study of US public and private schools (level not reported although it appears to involve a diversity) a simple set of models was reported for particular single determinants, with separate analyses conducted for each year group. An attractive physical environment was not significantly associated with the following outcomes: twelfth graders (age 17–18 years) being drunk in the past 30 days; eighth graders' (age 13–14 years) marijuana use in the past 12 months; eighth graders' alcohol use in school in the past year; and eighth graders' use of cigarettes in the past 4 weeks in school. However, this exposure was associated with reduced alcohol use in school in the past year among tenth graders (age 15–16 years) (coefficient=-0.350; $p<0.01$) and twelfth graders (coefficient=-0.350; $p<0.01$). Neglected physical environment was not significantly associated with twelfth graders' cigarette use in the past 30 days. The total number of unobservable and unsupervised places in and around school was not significantly associated with twelfth graders' marijuana use in the past 12 months, nor with twelfth graders' use of cigarettes in the past 4 weeks in school, but it was associated with twelfth graders' use of alcohol in school in the past year (coefficient=0.142; $p<0.01$), tenth graders' use of alcohol in school in the past year (coefficient=0.142; $p<0.01$) and tenth graders' use of marijuana and other illicit drugs in school in the past year (coefficient=0.100; $p<0.05$).

In a cross-sectional study of Dutch secondary schools, Monshouwer *et al.*⁶⁹ found a strong association between school type and episodic heavy drinking among students aged 12–16 years. When compared with young people in pre-vocational education schools, those in lower general secondary education were less likely to drink heavily (OR 0.69; 95% CI 0.57 to 0.83) and those in upper general secondary education were even less likely to do so (OR 0.40; 95% CI 0.32 to 0.51), and those in pre-university education were least likely to do so (OR 0.26; 95% CI 0.20 to 0.33). It is important to highlight, however, that, although school type was classed as a determinant, the type of school students were enrolled in was determined from tests taken at the end of primary schooling. Therefore, students were likely to be systematically different in different types of school, and so school effects were likely to be heavily confounded by individual-level factors. Whether or not a school permitted alcohol use at school on special occasions was not significantly associated with heavy drinking, and neither was a school's level of sanctioning if students were caught using alcohol at school.

In a cross-sectional study of secondary schools in Bavaria, Germany, Piontek *et al.*¹³² report that school type was not significantly associated with smoking in the previous 30 days among students aged 10–21 years. However, it should be noted that the school type coefficient seemed to indicate a misspecification of the model because only a single coefficient was reported for a categorical variable with four categories. Although a complete smoking ban for students at/around school was significantly associated with reduced smoking (OR 0.62; 95% CI 0.42 to 0.92), an extensive smoking ban for adults at/around school, student smoking being monitored regularly at/around school and a school smoking policy involving constructive sanctions were not associated with smoking in the previous 30 days.

The above results suggest no clear pattern, with the subgroup of better-adjusted studies reporting different levels of school effects than the less-well adjusted subgroup of studies. However, given the small number of studies in each subgroup and the variety of measures reported, it would be difficult to draw firm conclusions about variations in effect by subgroup.

Discussion

Narrative overview of findings

We included 42 reports of multilevel studies of school-level health effects (from 34 data sets) in which school-level factors were measured 'objectively' (i.e. not merely data from individuals from whom outcome data were collected). We confined our narrative synthesis of results to studies that adjusted for an array of key potential confounders (setting strict and less strict thresholds depending on the comprehensiveness of adjustment) and did not adjust for factors that might mediate school effects on health. Six studies met our stricter definition of being appropriately adjusted and a further four met our less strict definition. These intervention studies addressed some aspects of our fourth RQ more than others. In terms of school-level influences, the studies of value-added education are intended as indicators of teaching and of pastoral support; the studies of school policies can be regarded as examining both how schools are managed and how they instil discipline; and the studies of school campus areas and observable compared with unobservable places are examining aspects of physical environment. The studies generally reported on measures of health behaviours (such as substance use and fighting) rather than on health or 'health promotion' outcomes. Substance use was a predominant although not exclusive focus of the multilevel studies. These multilevel studies did not report subgroup analyses relevant to assessing health inequalities.

Reflecting an earlier review,⁴⁹ we found that ICCs for health outcomes, where reported, were generally low, with notable exceptions such as for verbal victimisation¹³⁸ and exercise,¹³⁵ indicating that most of the variation in outcome measures was due to variability at the student level as opposed to the school level.

We found consistent evidence from cross-sectional^{53,64,66} and longitudinal studies^{53,65} of middle schools in the USA and secondary schools in the UK that schools in which attainment was higher than would be expected from the social profile of students, and truancy was lower than expected, had lower rates of substance use. The Tobler *et al.* study⁵³ is particularly valuable not only in providing longitudinal evidence and examining an outcome concerning group fighting in addition to substance use, but also in examining the extent to which these effects appear to be relevant not only to predominantly white, middle-class English adolescents but also to low-income, ethnic minority young people in US cities. Gross rates of attainment and truancy were not associated with student health outcomes and the authors of these reports interpret this to mean that what matters is the value that schools add rather than gross rates of attainment and truancy because the latter merely reflects the characteristics of students at intake rather than what the school as an institution provides.

Selection bias is unlikely to explain these results because this would require schools with both high value-added scores and high rates of substance use, and conversely schools with low value-added scores and low substance use, to have differentially tended not to participate. Reverse causality is also unlikely given that two of the studies were longitudinal and all of the studies used year-on-year averages of attainment and truancy data so that the data on exposures and outcomes reflected the experiences of different students. However, confounding might explain these results if some unmeasured or incompletely measured student or area factor was associated both with value-added scores and with student health outcomes. As Aveyard *et al.*⁴⁸ point out, the most obvious source of confounding arises from the same families both discouraging their children from engaging in risk behaviours and sending their children to effective schools. However, they point out that published data did not enable parents to identify exactly what value schools added, and the schools with the best attainment were not the same as the schools adding the most value. The two longitudinal studies adjusted for baseline measures of the risk behaviours being examined. Although these baseline risk behaviours might themselves be influenced by school-level factors, these do not represent cases of overadjustment for factors on the causal pathway because of the timing of the assessments. Nonetheless, a more useful approach would be to adjust for students' engagement in risk behaviours prior to their entry into the school, as West *et al.*^{130,137} (in studies that were excluded because of their adjusting for variables we judged to be potential mediators of school effects) were able to do.

The authors suggest that authoritative schools provide better support and an appropriate level of control of students' behaviour, and that students in such schools are more likely to adopt pro-school identities and values and are less likely to invest in anti-school risk behaviours such as smoking. The authors suggest that their measures of value added are proxies for this institutional culture. We would suggest two caveats, however. First, Aveyard *et al.*⁶⁴ and Bisset *et al.*⁶⁶ use cross-sectional data and, therefore, it is difficult to be confident about causalities implied by associations. Second, these studies used routine data to develop relatively simple measures of these school-level factors and so cannot provide insight into what processes might underlie these apparent effects. Other studies might shed some light on this. As reported earlier, factors such as strong leadership, student involvement, high expectations and frequent evaluation and praise are reported to explain school differences in attainment.^{18–21}

Findings on the influence of school policies were mixed. A German cross-sectional study of secondary schools reported that a complete smoking ban for students at/around school was significantly associated with reduced smoking, although monitoring of students' smoking and sanction types for those caught were not associated with reduced smoking.¹³² However, a cross-sectional survey of secondary schools in the USA and Australia found no association between various forms of school smoking policies and any measures of student smoking.¹³¹ These differences may reflect a 'ceiling' effect for the impact of smoking bans, which have already been widely implemented across schools in the USA and Australia but not in Germany, where student smoking has traditionally been more tolerated by school authorities. Earlier studies in the USA and the UK not included in our narrative review because of methodological limitations have also suggested that school smoking policies are associated with reduced smoking.¹³⁶ The finding from a cross-sectional study of Dutch secondary schools that whether or not a school permitted alcohol use at school on special occasions, as well as a school's level of sanctioning if students were caught using alcohol at school, was not associated with heavy drinking among students is likely to reflect such a policy being unlikely to have much power given the limited number of special occasions when alcohol use would even be considered possible, such as school trips and parties.⁶⁹

These studies relied on cross-sectional data, which do not provide strong evidence of causality because of the possibility of reverse causality. However, this is unlikely in the case of these school-level measures of policy because of the stability of policies from year to year. However, as with the studies of value added discussed above, confounding by unmeasured student, neighbourhood or school factors is a potential source of bias in estimating the effects of school-level policies on smoking and other student health outcomes.

We also reviewed evidence from one cross-sectional study which reported that students in US middle schools with larger total campus and playground areas per student have higher rates of student accelerometer-measured physical activity during school hours.¹³³ Our review also found evidence from one cross-sectional study of US high school students that the number of unobservable and unsupervised places in and around school was associated with tenth and twelfth graders' use of alcohol in school and tenth graders' use of marijuana in school in the previous 12 months but not twelfth graders' overall use of marijuana in the past year nor twelfth graders' cigarette smoking in school in the previous month. An attractive school environment was associated with tenth and twelfth graders' use of alcohol in school in the past year but not with whether or not eighth graders used marijuana in the past year or used alcohol in school in the past year or smoked cigarettes in school in the past month, or whether or not twelfth graders were drunk in the past 30 days. The authors conclude that these results suggest that schools may be effective in addressing risk behaviours that occur inside school but may not have much impact on behaviours outside school.⁶⁸ Finally, Henry *et al.*⁶⁷ reported from a cross-sectional study that the following school-level factors were not associated with alcohol use among students aged 13–14 years in high, junior high and middle schools in rural USA: whether eighth graders are located within the same school as high school students or are in separate schools, school size and pupil–teacher ratio. Reverse causality is unlikely to explain the results of any of these studies because of the enduring nature of the school-level measures, but confounding by unmeasured individual or neighbourhood factors might account for the associations.

The only study that examined subgroup effects was that reported by Markham *et al.*,⁶⁵ which found that in secondary schools in the West Midlands, UK, there was no significant interaction of the school-level measure of value-added education with whether or not pupils were regular smokers at baseline.

Strengths and limitations

As discussed earlier in relation to intervention outcome evaluations, our review was exploratory in nature, aiming to map the possible range of health outcomes that school-level exposures might affect. We developed guidance in stage 1 about which measures would count and which would not, although in practice we found that we had to make judgements. We included measures of physical and non-physical violence and attitudes to violence but not measures of weapon carrying or non-violent delinquency such as graffiti; and measures of feeling safe at school but not successful transition to secondary school. Although these judgements can be challenged, we did at least apply them consistently.

We decided to focus only on those studies in which measures of the school environment drew on 'objective' data, that is, they were not derived from aggregates of self-reports from the same individuals (usually students) who also provided data on health outcomes. This was partly a pragmatic decision to limit the number of reports to review to a feasible, yet still ambitious, number. It was also informed by our view that, when data on exposures and outcomes are derived from the same individuals, any associations found might merely reflect unmeasured characteristics of individuals unrelated to the effects of the school environment. We recognise that the objective measures in question were themselves heterogeneous: researcher observations, for example, being very different from teacher reports. Such measures may be subject to information bias that differs from that found with student-derived data; for example, it may be that teacher-derived measures present a more positive picture of schools than would student-derived data. However, it is unlikely that such information bias would be differential with regard to school-level influences and so it is unlikely to have biased our findings. An alternative approach would have been to have included studies regardless of whether these used 'objective' or 'subjective' measures of school-level factors, but excluded all studies from our review (and not merely our narrative synthesis) that did not take what we judged to be an appropriate approach to adjustment for covariates.

It was sometimes difficult to judge whether a covariate was appropriate or inappropriate to adjust for. Some covariates were treated as appropriate because they were determined to be stable psychological traits, despite our thinking quite plausibly that these could be affected by schooling. These included impulsiveness in Grunseit *et al.*⁹⁷ and Kim⁸⁸ and locus of control in Kim.⁸¹ Varsity sports participation was also treated as appropriate as Leatherdale *et al.*⁹⁰ regarded that this might be a measure of intrinsic athletic ability that schools recognise but do not affect, although we again had doubts about this. Although open to challenge, decisions about these covariates were not critical in determining that these studies did not meet either our stricter or broader definitions of appropriately adjusting for covariates.

Our inclusion criteria were focused on studies of children aged 4–18 years. However, we included two multilevel studies in which the age range went up to slightly older than age 18 years, reflecting the secondary education system of that country. We included only studies in which the age range encompassed those below the age of 18 years and did not include any studies of postsecondary education.

The heterogeneity of the studies we included meant that we could not undertake statistical meta-analysis and were confined to narrative synthesis. We urge caution in the interpretation of our findings of our narrative synthesis of multilevel model studies, as with those of our intervention studies. Our review did not set specific primary and secondary outcomes. In narrating the findings of studies we have tried to give equal weight to significant and non-significant associations and comment in our summing up on the balance between these, but this is no substitute for testing hypotheses defined in terms of a small number of primary and secondary outcomes.

Chapter 10 Research question 5: qualitative studies

Research question

Through what processes might these school-level influences occur, examined using qualitative research?

Methods

Inclusion and exclusion criteria

A total of 194 references included in the evidence map were relevant to RQ5: qualitative studies. We included qualitative studies in the in-depth review that explored the processes by which school-level factors relating to how schools are led and managed, or how they teach, provide pastoral care to or discipline students; student attitudes to school or relations with teachers; and/or the school physical environment might influence student health.

Two reviewers (FJ and AH) developed and applied an initial set of exclusion criteria to reports coded as addressing RQ5 in the evidence map based on the full text. Pilot screening was conducted on a random sample of 10 reports to test and refine the criteria. All reports were then double-screened by the two reviewers. The reviewers discussed any discrepancies in screening until agreement was reached. In cases in which an agreement could not be reached, a third reviewer provided input. *Table 6* outlines the exclusion criteria for sift 1.

A second set of rating criteria was then applied to all reports included from sift 1 in order to limit the review to reports that provide findings that are conceptually rich enough to facilitate meta-ethnography. A scale of 'high', 'medium' and 'low' was used to rate each study's level of (1) conceptual richness (whether or not authors go beyond a description of the findings and interpret them to develop concepts, theories or metaphors), (2) relevance in terms of its research aims and (3) relevance of its findings to addressing our RQs. A total of 21 reports that scored 'high' or 'medium' across all three categories were included in the in-depth review.

Quality assessment

The criteria used for assessing methodological quality were built on EPPI-Centre health promotion reviews.⁸ The quality criteria addressed the rigour of sampling, data collection, data analysis, the extent to which the study findings are grounded in the data, whether or not the study privileges the perspectives of children and young people, the breadth of the findings and the depth of the findings. The complete quality assessment tool is available in *Appendix 5*. Based on the results of these criteria, reviewers judged the study overall in terms of reliability and trustworthiness of findings on a 'low', 'medium' and 'high' scale. Reports were not excluded based on these quality assessment ratings; instead, they were used to inform our interpretation of the findings. A round of pilot quality assessment was conducted by four reviewers (AH, CB, FJ and HW) on a sample of two reports to test the tool and ensure consistency in applying the criteria. All reports were then split between two reviewers (FJ and HW) and assessed independently. All items were checked by another reviewer (CB or AH) and any disagreements were resolved by discussion.

Data extraction

Reviewers summarised the qualitative study findings in their own words and extracted any themes or concepts that the study authors reported. Four reviewers extracted findings from a randomly selected sample of two reports (CB, AH, FJ and HW) to pilot the tool and ensure consistency in extraction. All other reports

TABLE 6 Exclusion criteria for qualitative reports

Exclusion criterion	Guidance
Exclude 1: map	This study is not relevant to our review. It should have been excluded at the mapping phase
Exclude 2: student health	Exclude reports that do not provide an account of how student health is affected by aspects of the school
Exclude 3: priority area	Exclude reports that do not report on the following aspects of the school: school type, physical environment, school organisation and management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers
Exclude 4: research design (qualitative study)	Exclude if this is not a qualitative study (e.g. not based on interviews, focus groups, ethnographic research)
Exclude 5: language	Study is not written in English
Include	Include this study if it passes all exclusion criteria

were split between two reviewers (FJ and HW) and completed independently. All items were checked by another reviewer (CB or AH) and any disagreements were resolved by discussion.

Synthesis

Meta-ethnography is an approach to research synthesis originally developed by Noblit and Hare.¹⁴⁴ They argue that meta-ethnographic enquiry is driven by the desire to develop interpretative explanations and understanding from multiple cases of a given phenomenon by utilising research that is 'grounded' in the experiences of participants (p. 12).¹⁴⁴ The approach has since been applied by qualitative systematic reviewers in the field of public health, particularly for questions related to patient experiences (e.g. see references¹⁴⁵⁻¹⁴⁷).

The purpose of this meta-ethnographic synthesis is to conduct an interpretive enquiry into the themes, concepts and metaphors set out by authors investigating the ways in which the school environment might influence young people's health. We developed a general interpretation grounded in the findings of separate studies of the process by which schools might influence health by reinterpreting meaning across individual qualitative studies. We used a four-phase synthesis approach adapted from Noblit and Hare's¹⁴⁴ phases of meta-ethnographic synthesis (pp. 26-9).

Step 1: understanding the themes and concepts and their relations within each study

The reviewers read and reread all of the reports to gain a detailed understanding of the findings, theories and concepts proposed. We uncovered the 'second order constructs' (authors' interpretation of the data) by identifying concepts developed in the reports. For example, one author developed the concept of 'unowned space' to refer to locations in the school environment that were violence prone.¹⁴⁸ 'Unowned space' was thus identified as a 'second order construct'. In reports in which concepts were not explicitly named, but instead described in the authors' narrative, the reviewers attempted to capture the authors' interpretations using a sentence or two. For example, we summarised one author's interpretation of the nature of violence in a school in the following way: violence as a functional tool for social control in the context of chronic poverty.¹⁴⁹ The concepts were identified by one reviewer (FJ) and then checked by another (AF) to ensure that they were comprehensive and accurately reflected the authors' interpretations.

To preserve the relationship between concepts *within* an individual study, a paragraph was written that described each study's perspective or argument about the way in which school-level effects occur. It tied together the concepts within the study in a narrative. This process was completed by one reviewer (FJ) and checked by another (AF) for accuracy and comprehensiveness. All studies were then grouped according to health topic. The groups of health topics include (1) aggressive behaviours (violence, bullying and

harassment), (2) substance use (drugs and tobacco use), (3) diet (eating habits), (4) sexual health and (5) rules for going to the toilet.

Step 2: how concepts from individual studies are related

The concepts from each individual study (identified in step 1) were listed on paper and closely related themes across different papers were grouped together resulting in a set of overarching themes. In other words, the reviewers conducted a thematic analysis of themes from individual studies. This process was completed for each health topic separately, resulting in a set of five overarching themes (corresponding to the five health topics). For example, from the studies related to aggressive behaviours, the following set of related themes were identified: conflict and violence in schools as a source of identity and social status;¹⁵⁰ 'tough identities' in violent contexts;¹⁵¹ violence as 'entertainment' at school;¹⁵² and drug use as an alternative source of identity and bonding.⁹¹⁻⁹³ These themes were grouped together under the overarching theme, *performance, collective identity and bonding*. Two reviewers grouped themes from individual studies by health topic (FJ and AF).

Step 3: developing the findings from the translations

The purpose of this step was to find the relationship between the overarching themes from each health topic (from step 2) and produce 'meta-themes' that attempt to capture all concepts across all health topics. Two reviewers (FJ and AF) discussed how themes from different health topics might relate to one another and how themes might differ. Ultimately, the reviewers produced a list of meta-themes by collapsing themes from different health topics and identifying unique ones. To draw out the findings under each meta-theme, studies that were rated 'high' in terms of quality and/or conceptual richness were chosen as 'index' papers from which we extracted findings; we then compared and contrasted these findings with the findings of a second study, and the resulting synthesis of these two studies was then contrasted with the findings of a third study and so forth. This process enabled us to develop a rich narrative that unpacks the concepts from individual studies as they relate to the meta-theme and ultimately responds to the review question.

Step 4: synthesising translations (from step 3)

From an interpretive reading of the translations across health topics (produced in step 3) we developed a general interpretation grounded in the findings of the included studies of the processes by which schools might influence health (Noblit and Hare¹⁴⁴ refer to this as the 'line of argument').

This was the ultimate aim of the reviewers as they completed all of the previous steps so that the synthesis was built up gradually through a cumulative process of interpretation and synthesis. One reviewer (FJ) who was most familiar with the data developed this explanation, which was checked for comprehensiveness, depth and accuracy by two other reviewers (CB and AF).

Overview of included reports

Flow of literature

In total, 194 qualitative reports examining the process by which schools might influence health were identified as potentially relevant in our map. The full texts of these reports were retrieved and screened in two sifts. In the first sift, 106 reports were excluded, the full texts of 22 reports were not available and eight reports were duplicates that were not picked up in earlier deduplication efforts (all studies were reported in English). The remaining 58 reports proceeded to the second sift. In this, 37 reports were excluded. The remaining 21 reports proceeded to the quality assessment and data extraction phase. Three of these references were 'linked' and thus counted as one study. Therefore, 21 reports of 19 studies were included in the in-depth review. A total of 10 studies examined aggressive behaviours, four studies examined substance use, two studies examined eating behaviours, two studies examined sexual health and one study examined the rules for going to the toilet. *Figure 11* provides a summary diagram of the flow of literature through the review.

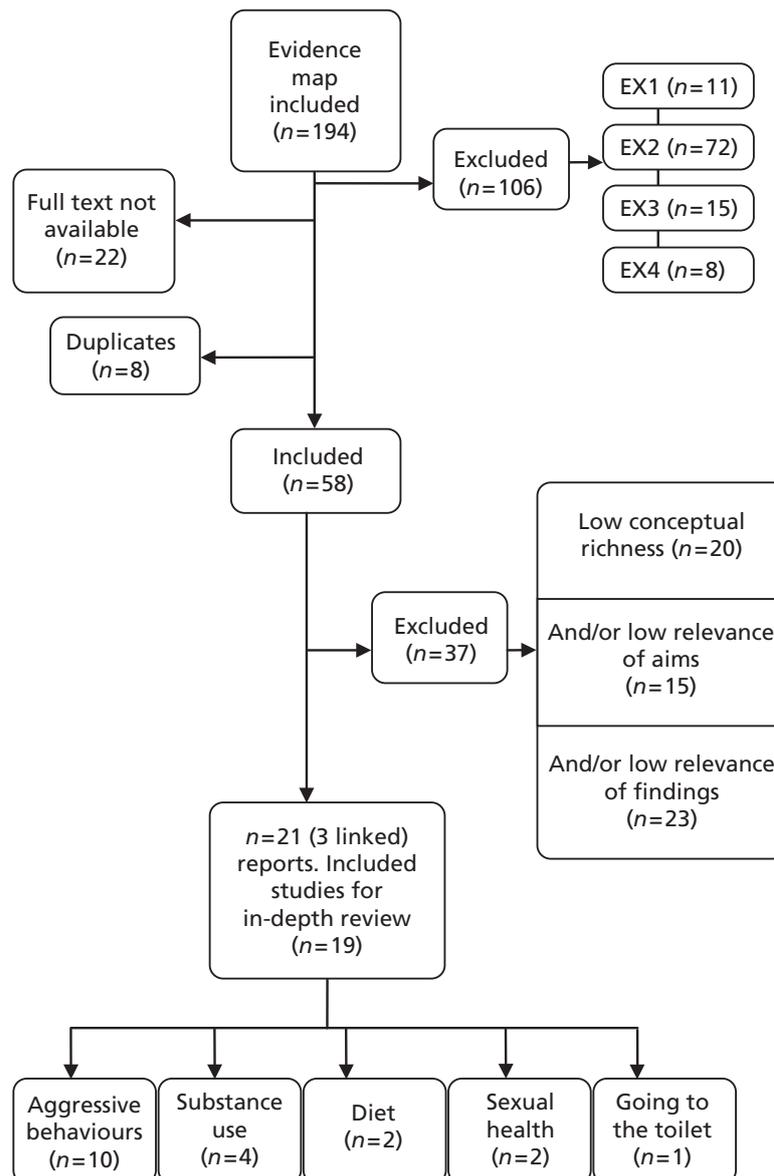


FIGURE 11 Flow of literature: qualitative studies (stage 2: in-depth synthesis).

Quality assessment

Most of the studies in this review ($n=14$) were considered to be of high^{91–93,148,150,152–154} or medium^{94,95,157–161} quality; five studies^{96,149,151,162,163} were considered to be of low quality based on the reliability and trustworthiness of the findings.

Three of the low-quality studies^{151,162,163} used ethnographic methods, such as participant observation, and were considered to be among the most conceptually rich. Reports were not excluded based on low-quality scores because reviewers felt that this could bias the review according to certain methodological approaches (e.g. interviews and focus groups rather than ethnographic approaches) and certain academic disciplines in which methods may be less transparently reported (e.g. anthropology). The evidence synthesis draws on a mixture of evidence from studies deemed to be of high, medium and low quality.

Study characteristics

Seventeen studies^{91–96,148–150,152,153,155–161,163} used one or more of the following methods to collect data: interviews, focus groups, participant diaries, school studies, open-ended questionnaires and/or general observation. Two studies, in the tradition of anthropology, used participant observation^{151,162} and two studies

used participatory techniques^{148,155}, in which researchers worked with students to map out different health behaviours using blueprints of the school environment and surrounding area, alongside interview and focus group techniques.

Of the 19 included studies, 10 were conducted in the USA,^{94–96,148,151–153,161,162} six in the UK,^{91–93,95,155,156,158,163} one in Australia,¹⁵⁹ one in South Africa¹⁴⁹ and one in Sweden.¹⁶⁰ In terms of health topics, eight of the US studies^{94,96,148,150–153,157,161,162} examined bullying, violence, intimidation, harassment and/or physical violence (hereafter described collectively as ‘aggressive behaviours’). The studies from Australia and South Africa also examined aggressive behaviours. The other two US studies examined sexual health⁹⁴ and drug use respectively.⁹⁶ Of the six UK studies,^{91–93,95,155,156,158,163} three examined tobacco, alcohol and/or drug use (hereafter described collectively as ‘substance use’),^{91–93,95,155} two examined students’ eating behaviours at school^{156,163} and one examined sexual health.¹⁵⁸ The study from Sweden examined rules for going to the toilet at school.

The majority of studies were conducted in high school/secondary school settings. Other types of schools examined were middle schools and primary/elementary schools. One study¹⁴⁹ was conducted in an environment of chronic poverty (South Africa) in a middle-income country. Of studies undertaken in high-income countries, nine studies^{95,96,150–153,156,162,163} were conducted in contexts of high levels of socioeconomic disadvantage; two studies were conducted with a high SES sample;^{94,159} and four studies explored students’ experiences across multiple schools and/or neighbourhoods that varied in terms of their socioeconomic characteristics.^{91–93,148,157,161} It is difficult to assess the socioeconomic context of three of the studies.^{155,158,160} The majority of studies were conducted in an urban context ($n=13$) and none was focused on a rural setting, although some studies include schools from a range of different urban, suburban and rural locations.

A range of different ethnic minority groups was included in these studies. African Americans were the most studied ethnic group (seven studies^{148,150–153,161,162} examined at least one school with a majority African American population). Some studies focused on other ethnic minority groups such as experiences of Latina/o young people in the USA⁹⁶ ($n=1$) and the experiences of ‘coloured’ young people in South Africa¹⁴⁹ ($n=1$). One study focused exclusively on the experiences of white young people.¹⁵⁹

In summary, the qualitative synthesis is largely informed by evidence from the USA and the UK. It will also have a bias towards the views and experiences of young people from disadvantaged communities and low SES families, young people from ethnic minority groups (particularly African American students) and those students who attend schools in urban settings. In terms of specific health behaviours, qualitative researchers exploring and theorising potential school effects have focused mainly on how schools might shape aggressive behaviours and substance use.

Results

Meta-ethnography of study findings

The first stage of the synthesis process involved grouping studies by their health topic and extracting the key concepts and themes from each individual paper. Second, reviewers conducted a thematic analysis of these themes and concepts across individual papers that addressed the same health topics. The outcome of this process is reported in *Figure 12* (see Step 2: overarching themes by health topic). Finally, we developed meta-themes, which were inductively arrived at by reducing the themes from across each health topic. The outcome of this process is also reported in *Figure 12* (see Step 3: meta-themes).

The findings are organised according to these meta-themes produced in step 3. This allows us to present a thick description of the key concepts that emerged and the inter-relationships between studies, comparing and contrasting across all health topics.

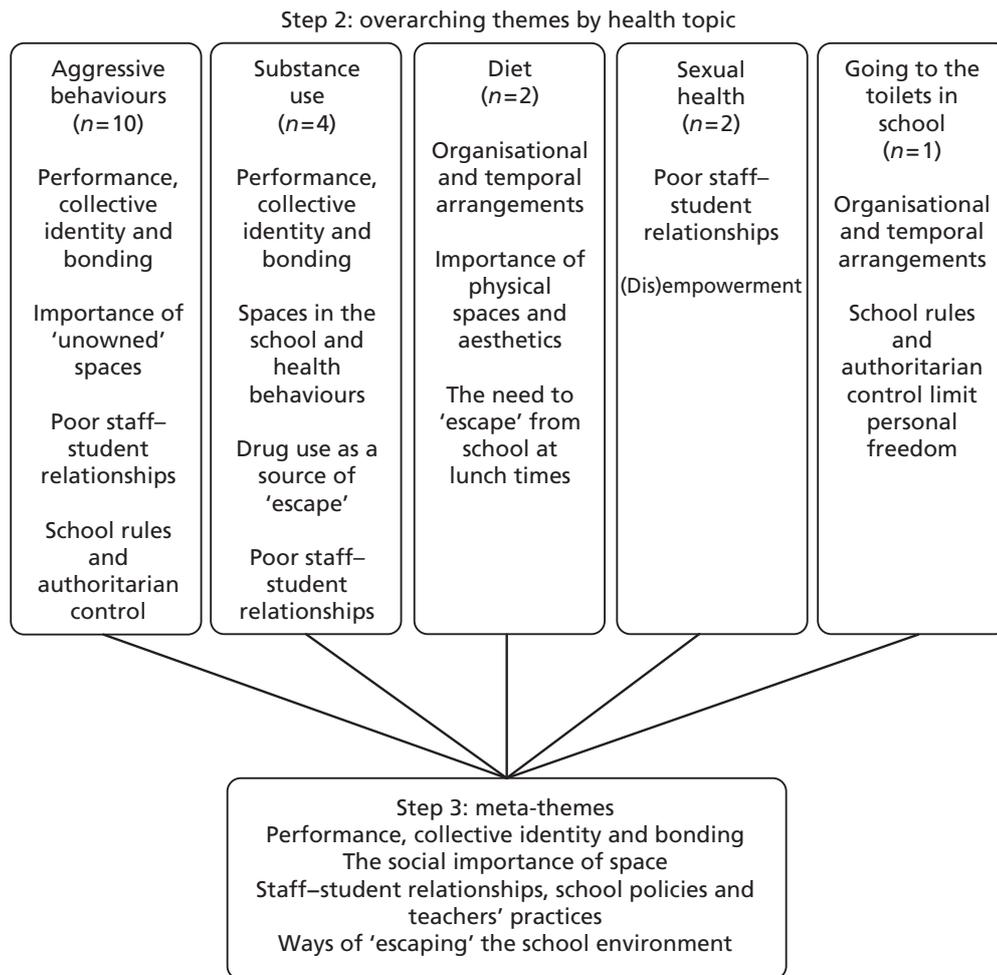


FIGURE 12 Arriving at meta-themes: reciprocal translation of included studies.

Meta-theme 1: performance, collective identity and bonding

The importance of 'toughness'

Young people are portrayed as often adopting 'tough' identities and 'street' styles at school based on aggressive behaviours, substance use and, in more extreme cases, carrying weapons and gang involvement. Through such *performances* young people can foster close relationships with 'tough' peers. These relationships appear essential for social support and 'back-up', especially in school environments that are considered to be unsafe (e.g. schools in low SES urban neighbourhoods). The importance of 'toughness' was a recurring theme across studies of both higher⁹² and lower¹⁶² quality conducted in the USA and UK:

On any given day one could observe small groups of boys leaning against the walls in the hall, posturing as I have seen them do on their street corners. Some wore dark glasses, black or brown high top Timberland boots, oversized jeans [. . .] a defiant, but cool, demeanour.

Ethnographic notes, USA, pp. 50–1¹⁶²

It's scary. You've gotta stand up for yourself – make your name, make friends. If you got friends then at least you know you're safe. Not being the person alone is important. You don't want to be one person alone.

Male student, UK, p. 245⁹²

The process of acting 'tough' appears to be an important source of status as well as social support,^{91–93,151,162} which helps students avoid being seen as weak and vulnerable in the school,¹⁵⁰ and this could be facilitated by using cannabis or adopting aggressive behaviours:

You smoke it [cannabis] for fun [but also] you wanna look bad. People think you're a bad boy or bad girl . . . with me they are cool and I'm safe with the boys here.

Female student, UK, p. 247⁹²

It's like a showcase here, you know, a lot of people, they just want people to watch 'em. . . [so] they try to make themself look hard.

Male student, USA, p. 197¹⁵⁰

A female student from a secondary school in a disadvantaged area of London, UK, described how 'geeky' students who chose not to adopt 'tough' identities felt unsafe and were vulnerable in schools:

I know for a fact some people don't feel safe. Like I know who them people are, they don't feel safe for sure [. . .] Most of them look like geeks 'cos they have glasses and they come in dressed like freaks from Camden Town and everyone just picks on them. [We] call them witches, and whatever.

Female student, UK, p. 246⁹²

In some cases, in studies of higher^{150,152} and lower¹⁶² quality, young people appear to go beyond 'tough' performances and actively engage in physical violence to showcase toughness. In this context, bystanders may encourage violent behaviour by cheering, creating a spectacle and space, or preventing others, including school staff, from intervening. This suggests that violent incidents in schools are often *group performances* through which the norms of acting 'tough' are collectively entrenched:

Charlie and his counterpart were throwing punches at each other, trying to push each other's head against the floor with all the strength that they could muster as they twisted their bodies together like twine. They were encircled by a ring of students locked arm-in-arm as they chanted in unison to the rhythm of the fighters.

Ethnographic notes, USA, p. 51¹⁶²

Student bystanders are reported to choose sides and, as such, physical fights in schools tend to make visible the tightly knit social networks within a student body.¹⁵² These behaviours again seem to reflect the social and cultural importance of establishing and maintaining respect in certain 'high-risk' school environments. In addition to the multiple potential impacts on students' health, these aggressive behaviours might reinforce poor education outcomes, teacher–student conflict and educational inequalities.

According to a lower-quality study by Cousins,¹⁶² the young black males in his study of an inner-city US high school tended to share 'tough' social and behavioural norms, which helped them to establish a sense of collective as well as individual identity. A higher-quality British study by Fletcher *et al.*⁹² similarly suggests that peer groups in low SES urban schools form around ethnic lines, which can shape students' health-related behaviours. For example, young black and dual-heritage men appeared to adopt cannabis as a strong source of identity to ensure their collective safety:

I'm not being racist or anything but I know for a fact that all black kids feel safe.

Female student, UK, p. 246⁹²

Although issues of race and ethnicity appear to be important for understanding how schools may influence students' health behaviours, this is not to say that certain groups of young people uncritically adopt behaviours such as drug use, but rather that the societal-level racial (and social class) context provides a 'tool kit' from which students can pick and choose behaviours.¹⁶² Cousins¹⁶² suggests that, if one considers the issues of black oppression and powerlessness at the societal or neighbourhood level, one might

understand violence in schools from a more socioculturally informed viewpoint. This is to say, although the school environment itself appears to be important, broader social and cultural influences also shape behaviours within schools.

Sex and 'tough' performances

The norms around showcasing toughness may also reflect the way in which the school environment maintains dominant masculine conventions, perhaps making young women particularly vulnerable. One author of a medium-quality study that adopted a feminist approach to examining the role of power and masculine entitlement in an Australian high school reported that:

The girls talked about how the peer group shaped particular boys' behaviours in ways that amplified masculine conventions and rigidified gender binaries. Here certain boys, according to the girls, made themselves 'feel better than everyone else' through 'dragging others down' and 'making them feel bad'. Such behaviours, particularly as they invariably escalated with an audience when things would be 'taken too far', were especially frustrating to the group of girls.

p. 6¹⁵⁹

Young women in this study reported that young men often subjected them to sexualised name calling and physical abuse (e.g. inappropriate touching, pulling hair and pulling skirts or trousers). The author suggests that, by doing this, young men assert their power and reproduce existing sexual inequalities in defining themselves in opposition to 'inferior girl-like femininity' (p. 6).¹⁵⁹ Teachers interviewed at a US high school within a high-quality study echoed these concerns regarding sexualised harassment and sexualised violence in schools:

One of our hall monitors saw a boy smack his girlfriend. And I said, 'You know, why would you do that?' And he said, 'Well she's gotta know I care about her!' . . . and he was serious.

Teacher, USA, p. 23¹⁴⁸

For some young women, their experiences of victimisation through these masculine performances could be 'kind of flattering' but in other ways also 'really horrible' (p. 8).¹⁵⁹ According to this medium-quality study,¹⁵⁹ this *double bind* that young women experience can be attributed to the notion that they feel a sense of legitimacy and power from ascribing to dominant masculine conventions.

Despite these strong masculine conventions, young women appear equally likely to be involved in violence and drug use in and around school. For example, a high-quality study by Astor *et al.*¹⁴⁸ unexpectedly found that over half of the violent events reported by high school students involved young women as both instigators and participants, and that young women were also involved in serious acts of aggression including stabbings, beatings and physical fights. This puts young women, particularly in low SES urban schools, at a *double risk of violence*: they are at risk of sexual harassment and assault from young men at their school and also put each other at risk. A high-quality study by Fletcher *et al.*⁹¹ suggests that the very vulnerable position of the most disadvantaged female students is actively reinforced by the ways in which the school environment is structured to support only the most 'perfect' and academically able female students.

Meta-theme 2: the social importance of space

'Unowned' spaces in the school

According to a number of high-quality studies, spaces that are unsupervised appear to be 'hotspots' for certain health-risk behaviours. For example, aggressive behaviours and substance use were often associated with areas such as hallways, staircases, toilets, changing rooms, empty classrooms or peripheral areas just outside the school.

Astor *et al.*^{148,153} used the terms 'unowned' and 'undefined' to refer to areas in the school that are largely unoccupied or unmonitored by adults, either at break times or throughout the whole school day. Spaces

appear to be 'unowned' because teachers are reportedly focused on *classroom-based instruction* and do not regard the supervision of the wider school environment as part of their professional responsibility. A study of five US high schools, in which the researchers distributed maps of the school and participants identified unsafe areas, found that all of the 166 reported violent events were in locations with few or no adults present, such as the parking lot, the lunchroom and hallways during transition times (p. 16).¹⁴⁸

In addition to an apparently increasingly narrow focus on classroom-based learning, school staff also reported a variety of other reasons for not intervening in student conflict beyond the classroom, such as the ambiguity of procedures, fear of harm and inadequate support systems (pp. 19–23).¹⁴⁸ These spaces thus seemed to become venues where students can engage in illicit or dangerous activities outside of the control of school authorities:

Most of the fights really kick off in the hallways.

Male student, USA, p. 198¹⁵⁰

the cafeteria is all the way on the top floor. . . and then the school is real big and. . . It's a whole lot of vacant parts of the school where you can just go. Aign't no teachers. . . just empty classrooms an stuff.

Male student, USA, p. 199¹⁵⁰

The idea of 'unowned' space might also be applied to understand the 'hotspot' locations for smoking. In a study comparing tobacco use in four high schools, students across all schools recognised a peripheral location either on or adjacent to the school property where students go to smoke with little fear of consequences (p. 1271).¹⁵⁵

In recent years, security guards, metal detectors and closed-circuit television (CCTV) cameras have been brought into schools as surveillance mechanisms in spaces with little or no adult supervision. However, these new security measures were often not considered to be effective or appropriate by students or staff:

All the cameras are gonna do is videotape, you know what I'm saying? They'll fight right in front of the camera too. . . some of them they'll be asking, 'Can I get that tape?

Male student, USA, p. 29¹⁴⁸

If it ever came to the point that we had to put metal detectors on our front doors and pass all kids through it, we've lost the battle. . . I think it's a message it sends. It's the impact it has on the total environment. You know, Big Brother watching.

School administrator, USA, p. 29¹⁴⁸

The study by Brunson and Miller¹⁵⁰ suggested how students' safety varied in different school contexts. Black males who had been expelled from mainstream schooling reported less violence at their alternative schools and perceived these to be much safer because secluded spaces were minimised and school sizes were small. For example:

There aign't no doors open that you can get in. The gym door stay locked and the other doors you can't get out or come in 'cept the front door.

Male student, USA, p. 199¹⁵⁰

It's so small you gotta try to get along with people, other gains. . . Ain't nobody here trip off [gang] stuff like that.

Male student, USA, p. 200¹⁵⁰

Astor *et al.*¹⁵³ also found that perceptions of safety in unsupervised spaces tended to vary by type of school and a school's social organisation. Middle schools were considered more unsafe by students than elementary schools because they were larger, with more 'undefined' space and greater overcrowding in hallways and

dining areas. In other words, *densely populated sites* and students interacting within *close proximity* were also perceived to be features of school environments that facilitated aggression.^{148,150} In low SES urban contexts, the dense grouping of unsupervised young people may provide enlarged audiences for showcasing tough identities, thus exacerbating existing problems.¹⁵⁰

Finally, the dynamics of school and neighbourhood violence tend to overlap in disadvantaged communities. In a US study exploring aggressive behaviours among high school students, the authors found that, when there was an incident of violence in the neighbourhood or a long-standing feud between gang-involved groups of young people in the community, it tended to fuel further conflict inside the school, with school time used to build up further 'hype' or settle disputes among students:

Well, say like the night before it might have been a shooting on somebody street, right. And this particular gang might think that this gang did it. And when they get to school, it's a lot of tension. Or it might be a fight at a mall or something, or at a store or something. And they bring it to school, you know.

Male student, USA, p. 202¹⁵⁰

Chaotic and unappealing spaces

Studies consistently reported that young people's eating habits were shaped by organisational and aesthetic features associated with school dining areas, and in particular it appears that opportunities for healthier eating may be constrained when schools' dining facilities are chaotic and aesthetically unappealing to young people. These processes also appeared to relate to broader problems associated with many schools' approaches to pastoral care and discipline.

In a high-quality study examining the eating habits of young people from disadvantaged backgrounds in Scotland, the authors reported that students disliked the organisational systems for break times and lunchtimes (pp. 460–1).¹⁵⁶ In particular, students disliked having more than one dining hall and different lunch sessions for different students; they also did not like that entry into the dining hall was organised by year group, whether they wanted hot or cold food and, in one school, according to whether or not they were wearing the correct school uniform. These regulatory constraints and social boundaries were described as highly frustrating because students often could not eat lunch with their friends and could not always access the foods they preferred. Young people also reported that their decisions about which food to purchase were largely dependent on the length of the queue at each counter as they did not want to use their limited break time queuing and with students they did not know. Dietary behaviours of young people were also reported as influenced by spatial factors in schools. For example, a lower-quality study that explored the food choices of young people in two schools with differing uptake of school meals found that students' attitudes and actions were influenced by the school's physical environment: whether or not this was an appealing 'social space'.¹⁶³ The school with better uptake of school meals had natural light coming through the dining room windows; chairs, doors and a salad bar with a matching colour scheme; and ramps near the serving counter so that smaller children could see the food. In contrast, the school with poor meal uptake was aesthetically unappealing with large curtains drawn and 'cheap moulded chairs'. The movement into the lunch room and movement within the lunch room tended to be more structured and less chaotic in the school with better uptake of school meals.

Everyone on the table waits for the queue at the counter to move along and then the teacher instigates the move up to the counter. 'Come on then' she says and she nods her head towards the queue. All the table rise together, picking up their plates from the table and make their way over to the queue. There is no rush to get to the queue first.

Ethnographic field notes, UK, p. 56¹⁶³

Another factor that seemed to influence lunchtime experiences for young people in this study was the presence of adults in the dining halls. Teachers tended to use lunch periods to prepare for afternoon lessons or have 'breathing space' away from the students. However, the way in which one school organised

afternoon lessons meant that teachers could eat with young people. Both students and teachers valued social time together in a non-academic setting:

Well, when I'm having lunch with the kids, if I'm sitting with them, erm, it's nice, cos I ask them what they're doing at the weekend or if they did anything nice at the weekend or if it's a Monday, erm, what their plans are. . . I just chat to them like, just a bit more friendly, I suppose. . . It's nice, cos they can chat about stuff and they can ask you things. They don't normally get that time, that's not me trying to pump them full of information, so, you know, and it's just a little bit more relaxed.

Teacher, UK, p. 57¹⁶³

According to a high-quality study, when lunch supervisors were the only school staff present their role was primarily to police specific spaces at lunchtime, and students felt unsafe, unsupported and uncomfortable, often eating quickly (if at all) to escape this environment. This is especially troubling when considering that the school canteen is a common site for fights because of the large number of students congregating in one place at lunchtime and the availability of ready-made props such as kitchen utensils and food to use as weapons (p. 198).¹⁵⁰ The presence of teachers in dining halls appeared to allow the lunchtime supervisors to concentrate on other aspects of their work such as cleaning, helping children to cut up their dinner and, crucially, encouraging children to select and eat the healthy food on offer (p. 57).¹⁵⁰

Students in another lower-quality study said that they experienced anxiety and difficulties with going to the toilet at break times in school because of concerns about their safety and other priorities such as wanting to spend time with friends.¹⁶⁰

Meta-theme 3: teacher–student relationships, school policies and teacher practices

The studies consistently suggest that having good relationships with school staff, particularly teachers, is integral to creating a healthy school environment. Students who have caring and supportive relationships with teachers report feeling safer in the school and appear more resilient towards participating in risky health behaviours, and more academically and socially engaged at their school. A good example of this was a low-quality study of pre-adolescent middle school students in the USA (mostly young women) which suggested that supportive teachers and school counsellors may foster resiliency in young people regarding drug use in schools.⁹⁶

However, staff–student relationships appear to be constrained and weaker when school rules are established and enforced by teachers without student input or consultation; teachers are 'out of touch' with the realities of young people's lives; there is limited supervision and support from teachers outside the classroom instructional environment; and teachers apply the rules inconsistently. In the following sections we explore how these factors impact on staff–student relationships and how, in turn, they relate to student health.

School rules that are established without student input or consultation

School rules are established and enforced by staff to maintain discipline in the school environment. However, according to studies of varying quality,^{91,155,160,161} this purely teacher-led approach may be counterproductive as students recognise the power imbalance and challenge the rules that they feel are unfair and disadvantage them because of their age or family background.

In one study of medium quality,¹⁶⁰ the rules regarding going to the toilet in a school stipulated that young people ought to raise their hand to report their need to use the facilities and request a hall pass. These rules, although developed to maintain order, were regarded as not recognising students' physical and developmental needs and as compromising their dignity:

It's a bit personal. . . you'd rather they didn't know what you're doing.

Student, Sweden, p. 221¹⁶⁰

The lack of student input and consultation over the rules was perceived as being due to a lack of respect from teachers.^{91,157} Although individual teachers may not actively disrespect students, the overall school structures for decision-making suggest a lack of respect. More generally, young people may find it difficult to accept being treated as passive and childlike in schools when they are forced to take on adult-like responsibilities outside of school. A quote from a female secondary school student from London, UK, illustrates this:

I've had to be an adult for, like, my whole life really but oh no, they just think they always know best 'cos they are the teacher and we are the students and we've gotta listen to them.

Female student, UK, p. 555⁹¹

In their high-quality study, Plano Clark *et al.*¹⁵⁵ suggest that schools should include students when designing policies. Student consultation and input in establishing and maintaining rules in the school is likely to facilitate a greater sense of ownership and connectedness to the school environment, ensuring that it is a safe and healthy place (p. 334).¹⁵⁷ A co-operative approach to rule-making might also foster a sense of mutual respect between students and teachers, social support and an environment more conducive for discussing sensitive issues.

Teachers ignoring the realities of young people's lives

Young people participating in the US and UK studies of both high^{91,92,148} and medium¹⁵⁹ quality consistently suggest that teachers are disconnected from the experiences of young black men and young women from disadvantaged families. This meant that school practices and policies were rarely engaging and meaningful for these young people. Young people cited feelings of 'hopelessness' and 'helplessness' regarding their life circumstances and this was often expressed through risk behaviours such as drug use or aggression:

I think, if you've got no hope, if you're surrounded by despair, then you don't see that following the rules, that good work and good deed will get you anywhere.

Female teacher, USA, p. 26¹⁴⁸

Most people who work [at school], they probably go in and it's all nice. They've probably got a mum and dad and a brother and a sister and dog and nice house, nice car [. . .] your background is what was you before this. Taking drugs makes you feel better.

Female student, UK, p. 558⁹¹

These examples illustrate that students from the most deprived backgrounds can make links between their low expectations for the future, isolating school experiences and use of drugs or aggressive behaviours to manage these issues.

In their medium-quality study, Gordon and Turner⁹⁵ also investigated the extent to which students model teachers' behaviours (exploring the concept of school staff as 'health exemplars') and also found that students and staff did not identify with each other. For example, school staff reported that, because of the significant age difference between them and the students, they could not understand or relate to student experiences. In contrast, students felt that staff did not understand them and had 'bad' attitudes towards them. This appeared to limit the extent to which young people viewed staff as being able to give a credible health message and support them.

The lack of teachers' supervision and support outside the classroom

According to studies of higher and lower quality, another key factor affecting student-staff relationships appears to be schools' narrow focus on teachers' responsibilities within the classroom environment, preparing students for public examinations and other forms of assessment, rather than across the whole school or in terms of pastoral support:

If there is a fight in the hall, there's no teacher there, and if there's a classroom right next to it. . . It's amazing how they don't hear it. When they come out, they just, like, stand there.

At-risk female student, USA, p. 329¹⁵⁷

Teachers in some settings appear to be increasingly focused narrowly on the cognitive (rather than social) dimensions of learning. Key disciplinary roles have often been outsourced to other professionals such as security guards or police officers, sometimes drawing on new technologies (e.g. CCTV cameras, metal detectors). However, guards are reported to overidentify with students and befriend a minority of them rather than ensure the safety of the whole school. Meanwhile, students are reported to dislike CCTV cameras, which can merely displace risk behaviours to new 'hotspots'. Some students are reported in high-quality¹⁵⁵ and low-quality¹⁵¹ studies to innovate in order to sneak guns or knives through metal detectors. In the worst case, guards were said to facilitate health-risk behaviours, such as drug use, or to have abused students:

Although the guards are discouraged by their superiors from 'fraternizing' with the students, they do often develop strong emotional relationships with them; we have known some guards who encourage students to study and to go to class; we have also known others who take drugs, sell drugs to students, have sex with them, and dispense favours.

Ethnographic field notes, USA, p. 176¹⁵¹

This 'outsourcing' of school discipline, most apparent in US inner-city contexts but not restricted to these settings alone, is therefore considered by young people to be ineffective in ensuring a safe school environment and appears to do more harm than good.^{92,148,153,155,157,161} Students consistently said in studies deemed to be of high¹⁴⁸ or medium¹⁵⁷ quality that teachers who are 'caring' or 'respectful' are the most effective in managing problem behaviours. Teachers characterised as 'caring' were said to be those who did not define their role too narrowly within the confines of their classroom. Such 'caring' staff were regarded as supporting better 'ownership' of public spaces in the school (see '*Unowned' spaces in the school*) and 'expressed a personal obligation or connection to the whole child regardless of the setting, location, time or expected professional role' (p. 24).¹⁴⁸

I would say that it is more like parenting. I talk to them [the students]. I don't keep my distance. I do not keep professionalism between us. I say what I really think, how I really feel. I break all of the rules.

Female teacher, USA, p. 25¹⁴⁸

Teachers basically fear the students. Teachers have abandoned the chore of acting as disciplinarians in the corridors, and the guards feel that if teachers did accept this responsibility, the students and the whole school would notice a significant difference.

Ethnographic field notes, USA, p. 176¹⁵¹

Although many teachers stated that they wished to be more involved with students, they cited that a lack of power made this difficult:

I can't make anything happen here. I have no power. . . There's nothing I can do. I have no voice.

Female teacher, USA, p. 25¹⁴⁸

Structural constraints at the school administration level, such as staff turnover and target-based education policies focused on academic attainment, were also implicated in the limited role of teachers in schools. In a medium-quality study of US suburban schools,¹⁶¹ staff reported that there was very little teacher time available outside of classroom learning because of high staff turnover. The 'gap' that emerges between teachers and students can also lead to student disengagement and greater risk-taking behaviours according to studies of high,^{93,148} medium⁵⁷ and low¹⁵¹ quality.

A medium-quality UK study¹⁵⁸ highlighted this connection between poor staff–student relationships, disengagement and teenage pregnancy. Many of the young women in this study who were pregnant while of school age reported that they disliked their secondary school experiences, which was often partly due to difficulties with learning, but mostly due to social factors such as poor relationships with teachers and bullying.

Teachers' inconsistent application of school rules

Another factor that appears to contribute to poor student–staff relationships is the inconsistent application of school rules, and this particularly relates to specific health-related behaviours such as bullying and smoking. In a medium-quality study of two suburban high schools (one with a mainly white intake and the other with a largely African American intake), Waldron¹⁶¹ found that rules and punishment were rarely put into practice in the way that they were presented in the school handbooks. Students frequently suggested that school staff were inconsistent in applying penalties. For example, teachers were said to apply less severe policies to children who were considered 'good' than to children who were considered 'bad' (p. 91), a distinction that students also regarded as being informed by social or economic class.

At the 'white' school, students considered 'bad' were white, working-class 'skaters' or 'punks' and, at the 'African American' school, 'bad' kids were 'hood-raised thugs' from lower-income backgrounds. Athletes were considered to be more of an asset to the school reputation in both schools and were therefore favoured, with less severe punishments for misconduct (p. 92).¹⁶¹ This might suggest that preconceived notions or stereotypes about misconduct or illicit activity among poorer young people may guide teachers' disciplinary treatments. It could potentially also suggest that teachers feel that they need to exert more social control on young people from poorer backgrounds. Feelings of unfairness experienced by young people in this context may result in a worsening of existing problems and greater conflict in the school.

The same study attributed the tendency of inconsistent application of school rules to the wider educational pressures placed on schools.¹⁶¹ Problems related to student misconduct were reportedly often kept 'hush-hush' by teachers in an effort to maintain the reputation of the school. It was perceived that problems could potentially sway public perceptions about the quality and safety of the school. Some teachers interviewed were concerned about this approach to discipline which disguised problems because it meant that genuine student concerns in the school were not addressed. Covering up or disregarding problems in the school meant that issues of student health may not be adequately addressed and that there is little likelihood of interventions being put in place to address poor health.

The unfair application of school rules between young men and young women may increase the vulnerability of young women towards sexual harassment and reproduce sexual inequalities in schools. For example, an Australian study deemed to be of medium quality¹⁵⁹ suggested that the reinforcement of sexual harassment in schools occurs through an over-reliance on prescriptive discipline systems, for example the discipline slip, which many young boys considered a reward rather than a punishment: 'I think they think they're cool because they got into trouble' (p. 11). Sexual harassment might also be reinforced through remedies such as giving a boy a basketball to play with, either to reward good behaviour or to distract them from bad behaviour:

But the thing is, they can't really do anything, so then they don't know what to do so then they just take them outside and talk to them or give them a basketball to play with and that's like rewarding them.

Female student, Australia, p. 11¹⁵⁹

This type of praise/reward was said to be problematic because it renders the girl invisible, delegitimises young women's positive school behaviour as insignificant and endorses misogynistic cultures in schools.¹⁵⁹ Furthermore, it may also create a risky health environment for young women as issues related to sexual harassment are not adequately addressed by teachers.

Inconsistent practices and unfair treatment may also operate at the level of the wider education system. One high-quality US study¹⁴⁸ reported that students from inner-city schools in one city reported that the school district would practice institutional discrimination, meaning that schools in deprived areas had reduced funding and resources and presumably less support for teachers to build relationships with students that can promote health.

Meta-theme 4: 'escaping' the school environment

Lunchtime as an opportunity to escape

Some students appeared to see lunch periods as a time to escape the school environment. Boys frequently reported playing football or other sports with friends during lunchtimes on the edge of the school site, as far away from the school 'instructional' environment as possible. Young women reported that lunchtime provided a relief, that they spent time 'hanging out' and were more likely than boys to escape the school site altogether and walk to local shops to buy food. Food was often eaten on the walk back to school or in local spaces surrounding the school that young people claimed as their own:

Just usually run to try and beat all the queues for the food [down the high street] and then like we go down to the wee pigeon bit [an area frequented by pigeons just outside the school grounds], sit, ate our lunch and then probably have a fag or two and then go back up the school.

Student, UK, p. 462¹⁵⁶

According to this high-quality study, for some students, going home for lunch appeared to be a way of temporarily escaping an environment that was considered unenjoyable or unfulfilling:

Usually your lunch only takes about ten, fifteen minutes to eat and there's usually nothing to do in the school so it's just we manage the time better when we just come home.

Student, UK, p. 462¹⁵⁶

The need to escape the school environment at lunch periods therefore has implications for young people's diet as they are less likely to purchase healthy foods, are more likely to visit fast-food restaurants and may not eat anything at all. It may also have implications for social relationships as some young people go home at lunch and therefore have little or no social contact with other students.

Substance use as a source of 'escape'

Using cannabis or other drugs was reported in studies of varying quality⁹¹⁻⁹³ as being a way of managing anxieties about school in the context of a lack of social support. Heavy drug use may also be a response to exam stress or a means to 'escape' the constant sense of academic failure:

If someone can't be bothered about school, like you're having a bad day then have a spliff in the morning and then it's a good day. Pressure and stress can make people take drugs. If people don't like the environment they're in they are not going to be comfortable and getting on at school.

Female student, UK, p. 247⁹²

Being 'stoned' also appeared to allow some underachieving students to have a credible excuse for their academic failure and an opportunity to escape the ongoing and public humiliation because of this. According to studies of high⁹¹ and medium¹⁵⁸ quality, British schools reportedly adopt increasingly authoritarian practices once students are preparing for public examinations (GCSEs) in years 10 and 11 (age 14–16 years), with more disruptive behaviours leading to significant repercussions such as long periods of isolation. This could lead some year 10/11 students to engage in more passive forms of resistance that were facilitated by getting 'stoned' before school:

When you're stoned you either just ignore the teachers or crack up!

Female student, UK, p. 557⁹¹

Using cannabis was used not only to escape the reality of the school environment but also to escape punishment from parents or difficult family situations. When students are experiencing a combination of problems at home and at school, heavy frequent drug use is most likely to be initiated as a form of self-medication.⁹²

Discussion

Summary of key findings

A total of 21 reports met our inclusion criteria and were included in the review. These addressed our fifth RQ about the processes through which school effects might occur. They focused on the aspects of the school environment most salient to and of most concern to students, for example schools' physical environments, teachers' lack of involvement in pastoral care and the outsourcing of discipline to security guards and new technologies. There was less focus on how schools are organised and managed and how teaching is delivered. In terms of outcomes, our studies focused mainly on violence and substance use with some attention to diet and to teenage pregnancy. There was surprisingly little data on emotional and mental health.

Our qualitative synthesis suggests the complex pathways through which poor school organisation and management may lead to physical and mental health harms. First, students' responses to schools in which they feel educationally marginalised or unsafe may actually exacerbate school disengagement and overall lack of safety. For example, when students adopt substance use and 'tough fronts'¹⁶⁴ based on aggressive behaviours to facilitate a 'safe' identity and build protective social support, this may reinforce conflict and unhappiness at school. Second, strong positive teacher–student relationships appear to be critical in promoting student well-being and avoidance of risk behaviour. Organisational features of secondary schools in the USA, the UK and elsewhere appear to undermine such relationships. In particular, through schools' limited involvement of students in decision-making, students can lose what social control theory defines as a 'stake' in their school community, thus increasing the likelihood that they will look for a sense of identity and social support through some of the risk behaviours suggested above.

The meta-ethnography also therefore resonates with the findings of the synthesis of the multilevel studies in suggesting that schools in which relationships between teachers and students are not strong are less likely to protect their students from harm and are more likely inadvertently to promote substance use and other risk behaviours. Qualitative research also provides new insights regarding how students' lack of satisfaction with school life may cause them to seek sources of 'escape', either through heavy patterns of substance use or by physically leaving school at lunchtime or for longer unauthorised spells. While the focus of the intervention and multilevel studies was on impacts on students overall, telling us very little about subgroup effects or health inequalities, a large proportion of the qualitative studies focused on young people growing up in socially and economically disadvantaged families and communities. This qualitative evidence also suggests that the extent to which schools protect and engage students, and ensure that teachers fulfil a broader role in young people's development, may be most important in such settings, because of the higher rates of violence, drug use and other risk behaviours in those communities, inequities in school funding and higher rates of staff turnover.

Strengths and limitations

The benefit of using a meta-ethnographic approach compared with narrative approaches is that it moves beyond providing a review of the accounts of individual studies and instead develops higher-order explanations. The inductive and interpretive approach helps illuminate meaning that can be valuable to policy-makers looking to understand behaviours in context, perceptions and unintended consequences. In this review the development of meta-themes can also be used to interrogate the theory synthesis and vice versa, building on established theories and identifying conceptual gaps in the literature. The value of a meta-ethnographic approach is supported by the remarkable consistency in the findings of studies of variable

quality undertaken in a wide range of settings, which differed by school system, deprivation level and ethnic make-up.

A potential limitation is that we may have covered some key concepts and themes in insufficient depth in attempting to translate the themes across studies and identify meta-themes. We attempted to preserve individual authors' interpretive frameworks by ensuring that all themes extracted from individual papers were accompanied by a narrative explaining how they are connected in order to report these relationships when synthesising the findings of multiple studies. Inevitably, however, some of this work is compromised when moving to higher levels of abstraction.

The findings presented in this chapter may not cover all of the relevant qualitative literature available. For example, we did not include reports that might address student health and the school environment when this is not the primary research aim of the study. Only those reports whose explicit and primary purpose was to address student health and its relationship with the school environment were included, which ignores a large body of work conducted by sociologists of education that may also help to understand the processes by which schools influence health outcomes.

Chapter 11 Discussion and conclusions

Review limitations

Deviations from protocol

We deviated from protocol in five minor ways. First, because of lack of time we decided not to search for research reports that cited already-included reports. Second, because of recruitment delays, we did not, in our initial sifting of identified references, have three reviewers working in parallel comparing answers in batches. Instead, sifting was carried out by CB working in parallel with each researcher as and when they started work on the project, still comparing answers in initial batches of at least three sets of 50 until at least 99% agreement (as against 95% in the protocol) was reached. Third, we did not extract report authors' conclusions about results because we concluded that these would not inform our review and synthesis. Fourth, our protocol stated that, for coding-included reports for the evidence map, two reviewers would code an initial sample in parallel to reach consensus and subsequent coding would be carried out solely by one reviewer. However, in practice, all coding was carried out in parallel by two researchers, meeting to resolve disagreements. Fifth, we consulted with young people twice near the beginning of the project to ensure that we examined areas of priority to them rather than, as originally intended, consulting with young people once on the evidence map and once on the draft final report.

In other cases we did not strictly depart from protocol but nonetheless changed our plans, which we report here in the interests of transparency. Our protocol is only explicit about not excluding reports by language of publication with reference to our initial mapping in stage 1, which we followed. However, we did exclude reports not written in English in our in-depth review in stage 2. This applied only to 12 studies examining school-level effects on student health. These non-English reports all had abstracts written in English and from these we judged that there was little chance of them having employed multilevel modelling to examine school effects. If these reports had been published in English we would have confirmed this by obtaining the full paper; however, we did not do this because of the time and expense that this would have incurred for what we judged to be little likelihood of benefit.

Our plans for in-depth review in stage 2 were to develop a narrative synthesis of all multilevel studies focused on 'objective' measures of the school environment (i.e. not derived from aggregations of data from the same individuals from whom outcome data were collected). We included a greater than expected number of reports in stage 2. Our review of study methods had established that, although many studies were poorly conducted and reported, there was nonetheless a critical mass of 10 reports that were well conducted and reported, and which took what we judged to be a sound approach to adjusting for covariates. We therefore decided to limit our narrative synthesis of multilevel studies to these 10 reports.

Other limitations

This was intentionally an exploratory review that examined a relatively heterogeneous group of interventions. An alternative approach would have been to use our mapping in stage 1 to determine a more homogeneous group of interventions for in-depth review. However, this would have gone against the conclusions from our preliminary summary of theory and stakeholder consultation that what was most useful was a synthesis focused on the health effects of schools' organisation, management and physical environment, and the way that schools provide teaching, pastoral care and discipline – which necessarily involved a diverse group of interventions and school-level influences. Nonetheless, we would argue that future reviews might now focus on particular aspects of schools' social and physical environments.

Although a strength of our review is that it aimed to synthesise intervention and school-level effects on a broad range of health outcomes, this review did not examine other social, psychological and

educational outcomes that might mediate effects on health. However, this would be an interesting focus for future work.

We focused our in-depth review on studies of student health, but not staff health. This was a pragmatic decision taken to ensure that we had a manageable number of reports. This is nonetheless an important area of research. One of the authors (CB) is currently engaged in a systematic review of organisational interventions to reduce teacher stress and promote well-being.

We took a very broad approach to searching electronic databases because of the multidisciplinary nature of the review and the variability in the keywords that are applied to these reports. This produced a very large number of reports to sift. It is likely that this process led to some pertinent reports being missed or accidentally excluded. We therefore complemented this approach in stage 2 by checking potentially relevant studies cited in those reports we had already included, as well as contacting authors of included reports for advice about any studies we had missed. However, we did not undertake any hand-searching of journals or web searches because the relevant literature was spread over a wide range of journals and other sources.

We coded references for our map based purely on title and abstract because of the time pressure created by the large number of reports we had to sift. Although this meant some relevant reports being excluded, as explained above, as well as some irrelevant reports being included, we nonetheless believe that the map is fit for its purpose, of describing the broad spread of research on the topic of the effects of schools and school environment interventions on health, and informing our stakeholder consultation.

Our selection of theories to inform our review hypotheses was partially subjective. We used multiple criteria to form a judgement about which theories to draw on. We did not simply require that a theory meet every criterion to be deemed primary because our judgements were necessarily more subtle than this. For example, the theory of human functioning and school organisation did not meet our criterion of simplicity. However, we judged that this theory was very strong regarding our other criteria and so included it. As a further example, several of our theories did not meet our quality criterion of addressing a range of health outcomes, but we judged them to be sufficiently useful in understanding school effects on ASB-related outcomes and so we opted to use them to inform secondary review hypotheses. We think that this balance between using clear criteria and making overall judgements is acceptable and appropriate given that these concerned the development rather than the testing of hypotheses.

Although we developed detailed guidance about which measures would 'count' as health outcomes, in practice we found that we had to make judgements. Although the wisdom of these judgements could be open to challenge, we did apply them consistently throughout our review, in relation to both outcome evaluations and multilevel studies of school effects on health. For example, some evaluations were excluded because what were described as health outcomes in fact involved composites that included health and educational variables together. We included conflict resolution practices but no other prosocial forms of behaviour such as sharing; measures of loneliness but not the quantity or quality of family or peer relationships; and measures of feeling safe at school but not engagement in education.

Our hypotheses did not prespecify which outcomes should be treated as primary and secondary. As a result, our narrative synthesis might run the risk of overinterpreting statistically significant findings, which may reflect the multiplicity of outcomes being reviewed rather than real intervention effects. However, to counter this we have tried within our narrative synthesis to give equal weight to findings of no significant statistical associations and in summing up the results to comment on the balance between significant and non-significant associations. Nonetheless, we recognise that this is no substitute for testing specific hypotheses defined in terms of a small number of primary and secondary outcomes and so we urge that our findings be interpreted cautiously.

Our synthesis of the theories informing empirical research in this area was highly innovative and useful in both determining the priorities for the in-depth review and interpreting empirical studies. However, it is clear from the preceding sections, which report implications of the empirical studies of school environment interventions and multilevel models of school-level health effects (see *Chapters 7–9*), that the empirical research cannot definitely determine the validity of our hypotheses, both because the empirical studies were not focused on testing these and because of limitations in the measures of school environment used in such studies and our focus on health outcomes, which meant that we could not examine the relevant evidence on the more proximal outcomes in the causal pathways linking the school environment to health.⁶⁰ We did not, for example, synthesise evidence examining how school environment interventions and school-level determinants affected student attendance, engagement or attainment outcomes.

It is also clear, as reported earlier, that the theories that we identified were biased towards the study of ASB. However, this simply reflected the theories that current empirical studies of the health effects of schools and school environment interventions are using. Although this might influence how our overall synthesis is presented in this chapter, it should not bias the presentation of the results of the constituent reviews presented in *Chapters 7–10* because these simply present study results and their implications for our RQs rather than aiming to assess our review hypotheses.

Key results

Stage 1: identifying and describing reports

Searching

A total of 82,775 references were retrieved from the database searching. After removing duplicates, 62,329 references were left to sift. In total, 1144 references were included in the evidence map.

Evidence and theory map and stakeholder consultation

Most of the 1144 references were from primary research studies, conducted mainly in high-income countries such as the USA, the UK, Australia and Canada. The main health topics identified at the mapping stage were violence, bullying, harassment, diet and physical activity. Management or school polices, catering services or vending machines, and sport or active transport were the three main aspects of the school environment that the references included in the map referred to in their titles and abstracts. Based on the evidence map and consultations with young people, teachers, policy-makers and researchers, we opted to focus in our stage 2 in-depth synthesis on schools' organisation and management, teaching, pastoral care, discipline and physical environment. We undertook a further stage of sifting to apply new exclusion criteria to identify the relevant reports for in-depth review.

Stage 2: in-depth synthesis

After applying further exclusion criteria, our in-depth review examined 82 separate reports: two stand-alone theory papers, 16 intervention outcome evaluation reports (10 studies), one additional report of a process evaluation (four studies in all), 42 reports of multilevel studies of school-level health effects (34 data sets) and 21 reports of qualitative studies of processes of how schools might influence health.

Theory synthesis

In total, 24 theories were cited either in two stand-alone theory papers or in reports of empirical studies addressing our other review questions. The most commonly cited theories were ecological systems theory (cited in 10 reports), social control theory (cited in $n=6$), social disorganisation theory ($n=5$), social learning theory ($n=4$), the theory of human functioning and school organisation ($n=5$) and social cognitive theory ($n=4$).

Three theories fully met and 10 theories partially met our key criterion to determine which theories to focus on, i.e. whether or not the theory suggested which specific aspects of the school institution might influence health.

Outcome evaluations

Research question 2 asked: What are the health effects of interventions aiming to modify how schools are organised and managed, how they teach, provide pastoral care to and discipline students, and/or the school physical environment? What are their direct and indirect costs?

Six RCTs (10 reports) and four quasi-experimental evaluation studies (six reports) were included, although the quality of these evaluations was generally quite poor, particularly in terms of attrition and adjustment for clustering in the analysis, and small sample sizes. These intervention reports addressed most aspects of our second RQ, for example the AAYP and HSE projects modified the way that schools were managed and addressed pastoral care; CDP changed the way that classroom teaching occurred; CAPSLE and PeaceBuilders modified school discipline; and the playground interventions modified school physical environments. The study reports examined measures of health (such as injuries and social anxiety), health behaviours (such as aggression and condom use) and 'health promotion' outcomes (such as attitudes to aggression); however, they provided few data relevant to health inequalities or costs.

Nine papers reporting on five evaluations (all but one report/study involving a RCT) examined interventions that encouraged staff and students to build a stronger sense of community and/or better interpersonal relations.^{44,54,55,58,60,119–121,123} Such studies have been conducted in a range of school settings, including elementary, middle and secondary/high schools, although all except the quasi-experimental evaluation of the HSE intervention (UK) were conducted in the USA. Studies reported a range of significant benefits for some but not all outcomes. Among US elementary school students, by grade four the CDP intervention was reported as having positive effects on three of the four measures of cognitive social problem-solving in conflict situations, as well as teacher-reported students' solving of minor interpersonal problems, but no effects for self-esteem.^{54,55,58} By grade six it had benefits for loneliness/social dissatisfaction and social anxiety, and reduced student self-reported loneliness/social dissatisfaction and social anxiety. The AAYP intervention was delivered in Chicago elementary and middle schools.⁴⁴ Compared with schools in the curriculum-only arm, schools receiving the school environment component plus curriculum saw non-significant beneficial trends for violence, recent sexual intercourse and condom use among boys at the 3-year follow-up in grade eight. Only for two non-health outcomes (school delinquency and an overall aggregate measure combining the health outcomes with school delinquency and provoking behaviours) were differences statistically significant. Among girls, there was a significant increase in self-reported condom use, but no other effects. The CAPSLE intervention reduced peer-reported victimisation and aggression, self-reported aggression and empathy and peer-reported aggressive bystanding, and increased peer-reported helpful bystanding among third to fifth graders in US elementary schools at 2 years' follow-up.¹²⁰ CAPSLE also reduced a decline in self-reported empathy. These benefits were not maintained at the 3-year follow-up other than for peer-reported helpful bystanding. There were no effects for self-reported victimisation and self-reported beliefs in the legitimacy of aggression. The PeaceBuilders intervention delivered in US elementary schools had significant effects on teacher-rated aggression among third to fifth graders, but not grade K–2 students; however, there were no intervention effects on child self-reported aggression.^{119,121} Effects for teacher-rated aggression were larger for students with higher baseline aggression, although no test for interaction was performed. At the 1-year follow-up there were intervention effects on injury-related and fighting-related visits to the school nurse. The HSE intervention was associated with significantly higher reported rates of feeling safe at school post intervention among year 7 students, and there were differences of borderline significance in self-reported teasing of others, hurting of others or been in a fight. There were no significant differences regarding self-reported measures of being teased or threatened weekly/hurt ever in this school, belief that will try illegal drugs or smoke a cigarette in the future or belief that will get drunk before age 16 years.⁶⁰

The strongest evaluation in this category, that by Flay *et al.*⁴⁴ of the AAYP intervention, reports the least promising results of school environment intervention. However, although not a focus of this review, compared with the no-intervention comparison for boys, both the school environment plus curriculum arm and the curriculum-only arm provided significant benefits. The AAYP investigators themselves concluded that the environment component may have been critical to the overall impact of the combined environment/curriculum intervention based on the substantially larger effect size of the combined intervention than of the curriculum-only intervention, each compared with the no-intervention comparison, for example with regard to violent behaviour (47% compared with 35% reduction), drug use (34% compared with 32%) and recent sexual intercourse (65% compared with 44%).

Taken together these studies suggest the potential of school environment interventions to bring about benefits, particularly regarding measures concerned with violence and aggression. There was no evidence of any interventions causing harms. However, most of these studies were subject to important methodological limitations, such as not adjusting for clustering and high attrition. The strongest study in this category, that on the AAYP intervention, reported the most equivocal results, although the authors interpreted these as suggesting the potential benefits of school environment intervention. We conclude that the evidence base for this category of school environment intervention is promising but not definitive.

Two RCTs each reported on by a single report assessed interventions that combined changes to American middle schools' food and physical activity environments alongside actions that aim to empower students to contribute to achieving these changes. These studies were relatively well conducted, although the evaluation of HYP was subject to high and differential attrition. Both studies reported intervention benefits for some measures of student physical activity but neither reported intervention benefits for any measures of student healthy eating. Although the evidence for the effectiveness of these interventions is therefore somewhat stronger than the evidence for interventions addressing sense of community and interpersonal relations, it is not possible to conclude whether or not such interventions to modify the school food/physical activity environments might have achieved the same effects even in the absence of empowered student participation. However, as reported above, mediation analysis in the HYP study did suggest that student empowerment may be important.

Three quasi-experimental evaluations reported in a total of five reports examined the effects of interventions to improve playgrounds in British elementary schools either by providing new physical structures or merely providing surface markings. These reported mixed findings regarding effects on students' physical activity. There were indications from one study that the benefits of improving playground physical structures may be greater for younger children and when recess was longer. The stronger of the two studies examining playground surface markings suggested that these were associated with significant effects immediately after intervention on both VPA and MVPA. Given the limitations of these three studies, we must conclude that there is not currently a strong evidence base for the effects of these interventions.

Thus, although not providing definitive evidence of the effectiveness of school environment interventions, the studies reviewed suggest that, overall, such interventions have the potential to promote students' health.

Seven of the reports of intervention outcome evaluations examine subgroup effects, but only one examined effects by SES. Furthermore, subgroup reporting was carried out inconsistently and three failed to report tests for interaction so these studies provide little information on the likely impact of school environment interventions on health inequalities. Only the AAYP intervention was delivered in an explicitly deprived area; this study supports the notion that school environment interventions addressing low-income communities are feasible. None of the evaluations report on cost-effectiveness. With the exception of the playground interventions reported by Ridgers *et al.*¹²⁴ and Stratton and Mullan,¹¹⁸ none of the studies reported on intervention costs. The descriptions of the interventions do not provide enough data to enable estimates of costs, particularly in terms of the costs for schools that arose as a result of their participation in the interventions.

Our focus is interventions that aim to modify the school environment without simultaneously addressing school health curricula. We have concluded that there is evidence for the potential of such action to promote student health, but that the evidence is not currently definitive. We will assess the implications of these results for our review hypotheses below.

These results should not be taken as suggesting that there is no definitive evidence for the effectiveness of health-promoting school interventions or co-ordinated school health programmes that, as well as addressing school ethos or environment, include curricular and community components. A Cochrane review of the effects of health-promoting school interventions comprising changes to school environment/ethos, curricula and parent/community engagement is proceeding in parallel to our own review and is due to report later in 2012. To develop a fuller picture of the effects of school environment interventions, the results of our own review should be read in conjunction with those of the Cochrane review.

Process evaluations

Research question 3 asked: How feasible and acceptable are the school environment interventions examined in studies addressing RQ2? How does context affect this, examined using process evaluations linked to outcome evaluations reported under RQ2?

Of the 16 included outcome evaluations, five reports included process evaluations and one further linked process evaluation paper was found by reference checking. These employed a range of research methods, most frequently drawing on quantitative data collected from students and/or teachers. These reports addressed some aspects of our third RQ more than others. Although most examined feasibility or fidelity in some way, fewer examined acceptability and only one study used a mix of quantitative and qualitative methods to examine local context and how this influenced intervention processes. Process evaluations reported largely positive results regarding intervention feasibility, fidelity, reach and acceptability, although differences in methods prevent any comparison of the delivery and uptake of each intervention. The single study that examined context suggested that this was important, facilitating implementation when this built on schools' existing ethos and when senior staff championed the intervention.

Multilevel studies

Research question 4 asked: What are the effects on health and health inequalities among school students aged 4–18 years of school-level measures of school organisation and management, teaching, pastoral care and discipline, student attitudes to school or relations with teachers, and/or the physical environment (measured through 'objective' data), examined through multilevel quantitative designs?

We included 42 reports of multilevel studies of school-level health effects (which drew on 34 data sets) in which school-level factors were measured 'objectively' (i.e. not merely data from individuals from whom outcome data were collected). We confined our narrative synthesis of results to studies that adjusted for an array of key potential confounders (setting a strict and less-strict threshold depending on the comprehensiveness of adjustment) and did not adjust for potential mediators of school effects on health. Six studies met our stricter definition of being appropriately adjusted and a further four met our less-strict definition.

We found consistent evidence from cross-sectional^{53,64,66} and longitudinal^{53,65} studies of middle schools in the USA and secondary schools in the UK that schools in which attainment is higher than would be expected from the social profile of students and truancy is lower than would be expected have lower rates of substance use. The study by Tobler *et al.*⁵³ is particularly valuable not only in providing additional longitudinal evidence and examining a group-fighting outcome in addition to substance use, but also in examining the extent to which these effects appear to be relevant not only to predominantly white, middle-class British adolescents but also to low-income, ethnic minority young people in US cities. However, confounding from some unmeasured or incompletely measured student, neighbourhood or area factor might explain these results. These studies also used routine data to develop relatively simple measures of these school-level factors and so give little insight into what processes might underlie these apparent effects.

Findings on the influence of school policies were mixed. A German cross-sectional study of secondary schools reported that a complete smoking ban for students at/around school was significantly associated with reduced smoking¹³² (although monitoring of students' smoking and sanction types for those caught were not apparently influential), but a cross-sectional survey of secondary schools in the USA and Australia found no association between various forms of school smoking policies and any measures of student smoking.¹³¹ These differences may reflect a 'ceiling' effect for the impact of smoking bans, which have already been implemented across schools in the USA and Australia but not in Germany. A cross-sectional study of Dutch secondary schools reported that whether or not schools permitted alcohol use at school on special occasions was not significantly associated with heavy drinking among students age 12–16 years and neither was a school's level of sanctioning if students were caught using alcohol at school. However, this finding is likely to reflect the fact that such a policy would be unlikely to have much traction given the limited number of special occasions when alcohol use would even be considered possible at school.

A cross-sectional study found that students in US middle schools with larger total campus and playground areas per student have higher rates of student accelerometer-measured physical activity during school hours.¹³³ A cross-sectional study of US high school students found that the number of unobservable and unsupervised places in and around school was associated with tenth and twelfth graders' use of alcohol in school and tenth graders' use of marijuana in school in the previous 12 months but not twelfth graders' overall use of marijuana in the past year nor twelfth graders cigarette smoking in school in the previous month.⁶⁸ An attractive school environment was associated with tenth and twelfth graders' use of alcohol in school in the past year but not with whether eighth graders used marijuana in the past year, used alcohol in school in the past year or smoked cigarettes in school in the past month, or whether or not twelfth graders were drunk in the past 30 days. The authors conclude that these results suggest that schools may be effective in addressing risk behaviours that occur inside school and among older students but may not have much impact on behaviours outside school.

Finally, Henry *et al.*⁶⁷ reported that the following school-level factors were not associated with alcohol use among students age 13–14 years in high, junior high and middle schools in rural USA: whether eighth graders are located within the same school as high school students or are in separate schools, school size and pupil–teacher ratio.

Multilevel studies of school effects tell us very little about the impact of schools on health inequalities. The only well-adjusted multilevel study of school effects that examined subgroup effects (by baseline health behaviour not SES) was that reported by Markham *et al.*,⁶⁵ which found that in secondary schools in the West Midlands, UK, there was no significant interaction of the school-level measure of value-added education with whether or not pupils were regular smokers at baseline.

These studies relied on cross-sectional data but reverse causality is unlikely because of the stability of policies from year to year. However, as with the studies of value-added education, confounding by unmeasured student, neighbourhood or school factors is a potential source of bias.

Qualitative studies

Research question 5 asked: Through what processes might these school-level influences occur, examined using qualitative research?

A total of 21 reports from 19 studies met our inclusion criteria. The studies reported on suggest the complex pathways through which poor school organisation and management may lead to physical and mental health harms. First, students' responses to schools in which they feel educationally marginalised or unsafe may actually exacerbate school disengagement and the overall lack of safety in the school. Second, strong positive teacher–student relationships appear to be critical in promoting student well-being and avoidance of risk behaviour, but organisational features of secondary schools in the USA, the UK and elsewhere appear actively to undermine such relationships. Students can thus lose what social control theory defines as a

'stake' in their school community, increasing the likelihood that they will look for a sense of identity and social support through risk behaviours such as substance use and violence.

The meta-ethnography suggests that schools in which relationships between teachers and students are not strong are less likely to protect their students from harm and more likely inadvertently to promote substance use and other risk behaviours. Qualitative research also provides new insights regarding how students' lack of satisfaction with school life can cause them to seek sources of 'escape', either through heavy patterns of substance use or by physically leaving school, either at lunchtime or for longer unauthorised spells. Qualitative studies commonly focused on young people growing up in socially and economically disadvantaged families and communities. This qualitative evidence also suggests that the extent to which schools protect and engage students, and ensure that teachers fulfil a broader role in young people's development, may be most important in such settings because of the higher rates of violence, drug use and other risk behaviours in those communities, inequities in school funding and higher rates of staff turnover.

Overall synthesis

Consultation with policy and practice stakeholders

We consulted with individuals working in policy ($n=2$), research ($n=1$) and practice ($n=1$) on a first draft of this report, which at that point summarised findings from each of the different sections of our review but lacked an overall synthesis. This consultation aimed to get advice regarding what might be our key messages and how these might be disseminated in order to influence policy and practice. These discussions supported our aim of harnessing the evidence from the different sections to consider and refine existing theories and to ensure that any recommendations for interventions we produced stressed the importance that these be rigorously evaluated before any wider implementation.

Implications of our evidence for the primary and secondary review hypotheses

Our review of theories and stakeholder consultation in stage 1 led us to focus on how schools are organised and managed, are designed and built, and provide teaching, pastoral care and discipline. These are more 'upstream' determinants of student health than determinants such as what food schools provide or how they deliver PE. We then reviewed, in depth, theories about how schools influence student health cited in stand-alone papers or in empirical studies included in our review. We identified three primary and 10 secondary theories (*Table 7*) that were most relevant to understanding what specific aspects of schools affect health. We will now summarise what the empirical studies we reviewed suggest about the validity of these theories.

The intervention studies themselves did not set out to test any of our review hypotheses and were not informed by the theories from which these arose. Most of the interventions also employed multiple components so that it is not possible to take evidence of intervention effects as evidence of causal relations between the constructs within our review hypotheses and student health outcomes. As we will describe below, it is not clear, for example, whether the CDP's positive effects reflected its encouragement of student participation (which would support the social development model), its development of strong relationships between staff and students (which might support either social capital theory or the theory of human functioning and school organisation) or the process of giving students greater ownership – and thus a greater 'stake' – in their school (which would support social control theory). For these reasons, the intervention studies cannot offer very definitive evidence for or against our review hypotheses. Nevertheless, the possible ways in which the intervention studies might support or undermine our review hypotheses are considered below.

How the interventions addressed the domains specified in our primary and secondary review hypotheses is outlined in *Table 8*.

TABLE 7 Testable hypotheses of included studies

Theory	Hypothesis
Primary theories	
Social capital theory	Schools foster health by having a stable student and staff body, good relationships between staff and students, and a positive school ethos of stable, shared norms
Social development model	Schools reduce ASB by providing opportunities for students to participate fully in learning and community life and develop the skills necessary for such participation and ultimately enabling students to gain recognition
Theory of human functioning and school organisation	Schools foster student autonomy and health by reducing social boundaries between staff and students and among students, and ensuring student-centred framing of learning, management and other school systems
Secondary theories	
Deterrence theory	Schools reduce ASB by setting certain, severe and rapid punishments
Theory of reasoned action	
Theory of planned behaviour	
Ecological model of co-ordinated school health programmes	Schools foster health by promoting a supportive psychosocial environment, good safety facilities and opportunities/requirements for physical activity within the school
Integrated perspective on delinquent behaviour	Schools reduce ASB by ensuring that all students experience success in school activities
Moral authority theory	Schools reduce ASB by inculcating respect and not necessarily setting severe punishments
Problem behaviour theory	Schools reduce ASB by ensuring that students' educational and social problems are addressed
Social control theory	Schools reduce ASB by giving students some stake in the school community, perhaps by increasing student participation in decisions
Strain theory	Schools reduce ASB by ensuring that students can achieve their broader goals through school activities
Theory of triadic influence	Schools foster health by providing health education, reducing students' opportunities for engaging in risk, setting rules/norms against risky behaviours, enabling bonding between staff and students and providing good general education

Three RCTs^{44,54,55,58,120,121,123} and one quasi-experimental study⁶⁰ evaluated interventions that encouraged staff and students in a range of elementary/middle schools in the USA and in secondary schools in the UK to develop school climates characterised by a stronger sense of community. For example, the HSE intervention focused on developing better relationships; the CAPSLE intervention aimed to develop student participation in mediating conflict; and the PeaceBuilders and CDP interventions were aimed at enabling better interpersonal relations. These evaluations generally reported benefits. For example, CDP evaluations reported benefits for emotional health and conflict resolution.^{54,55,58,123} The CAPSLE and PeaceBuilders findings suggested positive effects on various measures of victimisation and aggression^{119–121} and the HSE study reported benefits for perceived student safety.⁶⁰ However, the AAYP evaluation reported fewer significant benefits regarding violence, recent sexual intercourse and condom use when comparing school environment change plus a social skills curriculum component with a curriculum-only intervention.⁴⁴ In contrast, although not a focus for our review, when comparing the combined intervention and the curriculum-only intervention with no intervention, findings for the combined intervention were more positive, suggesting that the school environment component may have been a key active ingredient.

TABLE 8 Theory domains addressed by intervention studies

Domain	Theory featuring this domain	Intervention addressing this domain
Student participation/stake in school community	Social development model Social control theory	CDP, AAYP, CAPSLE, HYP, M-SPAN
Student relationships with staff/other students	Social capital theory Theory of human functioning and school organisation Theory of triadic influence	CDP, AAYP, CAPSLE, PeaceBuilders, HYP, M-SPAN
Student-centred activities	Theory of human functioning and school organisation	CAPSLE, AAYP, HYP, M-SPAN
Students' cognitive development	Theory of triadic influence	CAPSLE, PeaceBuilders
Promotion of prosocial norms	Theory of triadic influence	CDP, AAYP, CAPSLE
Minimising students' exposure to risk	Theory of triadic influence	CAPSLE, PeaceBuilders
Sanctions against risk behaviours	Deterrence theory Theory of planned behaviour Theory of reasoned action	None
Schools' general psychosocial environment	Ecological model of co-ordinated school health programmes	CDP, AAYP, CAPSLE, PeaceBuilders, HYP, M-SPAN, playground interventions

With the caveats noted above, this evidence might lend broad support to each of our three primary hypotheses arising from the social development model (regarding the importance for health of participation in school activities), social capital theory (regarding the effects of trusting relationships) and the theory of human functioning and school organisation (regarding the importance of eroding rigid social boundaries between staff and students and that more student-centred framing of activities will enable better health outcomes). In terms of our secondary hypotheses, these studies suggest the importance of cognitive development, bonding between students, prosocial norms and protecting students from involvement in risk behaviour, which are key concepts in the theory of triadic influence. The importance of giving students a stake in their school community also suggests the possible validity of the hypothesis arising from social control theory. These findings from interventions that did not aim to increase the severity of sanctions accorded to aggression and violence but which nonetheless appeared to improve conflict resolution provide no support for our secondary hypothesis derived from deterrence theory. Finally, these interventions' broad focus on improving the general psychosocial environment also supports the hypothesis from the ecological model of co-ordinated school health programmes. A final caveat is that these studies of interventions aiming to build the school community generally had methodological problems such as small sample sizes and high or differential attrition.

Two RCTs^{63,122} focus on interventions in middle schools in the USA aimed at promoting healthy eating and physical activity through greater student participation in decision-making and modifying food and physical activity environments at school. Both reported mixed findings, but more positive effects for physical activity than for healthy eating. It is difficult to assess whether or not these effects might be attributed to student participation, although a mediation analysis conducted in one study suggested that this might be the case. With this caveat, the reported benefits for physical activity offer support to our review hypotheses about the importance of participation from the social development model and the theory of human functioning and school organisation. Three quasi-experimental evaluations^{117,118,124-126} examine changes to playground physical environments, but report mixed findings and do not resonate with constructs in our review hypotheses.

The multilevel studies provide greater insights regarding our review hypotheses because some of these explicitly set out to test hypotheses arising from the theory of human functioning and school organisation and because other studies, although not explicitly focused on our own review hypotheses, examined the specific effects of school-level factors that are closely aligned with the constructs in some of our hypotheses. How the multilevel studies examined the domains specified in our primary and secondary review hypotheses is outlined in *Table 9*.

TABLE 9 Theory domains examined by multilevel studies

Domain	Theory featuring this domain	Multilevel study examining this domain
Ensuring student commitment to instructional and regulatory orders	Theory of human functioning and school organisation	Aveyard <i>et al.</i> 2004; ⁶⁴ Bisset <i>et al.</i> 2007; ⁶⁶ Markham <i>et al.</i> 2008; ⁶⁵ Tobler <i>et al.</i> 2011 ⁵³
Setting sanctions against risk behaviours	Deterrence theory Theory of reasoned action Theory of planned behaviour	Evans-Whipp <i>et al.</i> 2010; ¹³¹ Piontek <i>et al.</i> 2008; ¹³² Monshouwer <i>et al.</i> 2007 ⁶⁹
Enabling physical activity	Ecological model of co-ordinated school health programmes	Cradock <i>et al.</i> 2007 ¹³³
Minimising students' exposure to risk	Theory of triadic influence	Kumar <i>et al.</i> 2008 ⁶⁸

In particular, four multilevel studies^{53,64–66} provide insights into the theory of human functioning and school organisation in relation to secondary/middle schools in the UK and USA, particularly with regard to substance use. A cross-sectional study by Aveyard *et al.*⁶⁴ of students aged 11–16 years in English secondary schools reports an association between more 'authoritative' schools (based on evidence that these schools had higher than expected attainment and lower than expected truancy) and lower rates of regular smoking. Drawing on the same data, Bisset *et al.*⁶⁶ find that authoritative school status is also associated with lower rates of various measures of alcohol and drug use. Markham *et al.*⁶⁵ use longitudinal data tracking students aged 13–14 years for 2 years in English secondary schools and report an association of borderline significance between authoritative school status and weekly smoking. Tobler *et al.*⁵³ also report on longitudinal data tracking students aged 11–12 years for 2 years in middle schools in the USA, using similar measures to those of Aveyard *et al.*⁶⁴ and Markham *et al.*,⁶⁵ and note that authoritative schools also had lower rates of smoking, drinking and drug use as well as group fighting in this context.

These findings are offered by the authors themselves as evidence in support of their theory of human functioning and school organisation. They suggest that authoritative schools provide better social support and an appropriate level of control of students' behaviour so that students in such schools are more likely to adopt pro-school identities and values, and less likely to invest in anti-school risk behaviours such as smoking. They suggest that their measures of whether or not schools had higher than expected attainment and lower than expected truancy based on their intakes (i.e. a school's 'value-added' score) are proxies for this institutional culture. These studies incidentally also support the hypotheses emerging from the social development model (i.e. that schools providing students with good opportunities to participate in learning and community life are more likely to enable the development of prosocial behaviours). Although we would agree that these studies provide support to the theory of human functioning and school organisation, we reiterate the caveats listed in *Chapter 9* concerning the cross-sectional nature of two of the four studies and the limited insights that such crude measures of school culture can provide into the causal pathways involved.

Evidence from the multilevel models we have reviewed is less well oriented towards testing the secondary hypothesis derived from deterrence theory that schools will limit risk behaviours by setting severe rules

against such activities, and currently provides no clear evidence for or against. Evans-Whipp *et al.*¹³¹ reported cross-sectional data for students aged 11–16 years at secondary schools in the USA and Australia, suggesting no association between school rules and policies on student smoking and student self-reported smoking outcomes (although they do find an association between strict enforcement of smoking policy and reduced perception of other students smoking on school grounds). Piontek *et al.*¹³² report no associations in cross-sectional data between smoking among students aged 10–21 years and restrictions on adult smoking, monitoring of student smoking and smoking policies in German secondary schools, but they do find an association between a complete student smoking ban and reduced rates of student smoking. As discussed earlier, these differences may reflect a ceiling effect: while German schools have often tolerated smoking among staff and older students so that smoking bans have scope for impact, in the USA and Australia there is a more pervasive intolerance of smoking so that the effects of differences in smoking policies between schools are difficult to detect. Cross-sectional data from Dutch secondary schools suggest that neither whether or not a school permitted alcohol use at school on special occasions nor the severity of sanctions was associated with student alcohol outcomes.⁶⁹

Cradock *et al.*¹³³ reported cross-sectional associations between the total area of the campus and playground of US middle schools and higher levels of physical activity during school hours, providing some support for our secondary hypothesis derived from the ecological model of co-ordinated school health programmes that schools foster health by enabling physical activity within school. The finding of Kumar *et al.*⁶⁸ that the total number of observable and unsupervised places in US high schools was associated with twelfth graders' use of alcohol in school in the past year, tenth graders' use of alcohol in school in the past year and tenth graders' use of marijuana and other illicit drugs in school in the past year but not with measures of substance use occurring both inside and outside school provides some support for our secondary hypothesis that schools can promote health by reducing opportunities for risk behaviour derived from the theory of triadic influence.

Our meta-ethnography of qualitative evidence enables us to develop a refined theory of how the school environment influences health.

Markham and Aveyard's⁶² theory of human functioning and school organisation suggests that schools comprise two 'orders': the 'instructional' order (which provides students with knowledge and skills) and the 'regulatory' order (concerned with students' conduct, character and membership of the school community). Our meta-ethnography suggests that there are two distinct, but overlapping, 'systems' that operate in the school environment: the student peer system (comprising student-led structures and processes) and the school institutional system (comprising structures and processes involving school management, teachers and other school staff and their use of technologies such as CCTV). Both systems are recognised in Markham and Aveyard's⁶² theory of human functioning and school organisation, but the former student-led system is not greatly elaborated on by them. Markham and Aveyard⁶² would refer to the separation of the student and institutional systems in terms of Bernstein's¹⁰⁵ notion of 'classification', which refers to the way in which boundaries between the students and the school staff are formed and modified over time. The strength of the boundaries between the student population and the school's other institutional actors is theorised as determining student capacity for affiliation and practical reasoning and, ultimately, their health outcomes. For example, students who remain uncommitted to the instructional and/or regulatory orders are theorised as more likely to become instead committed to anti-school peer groups and behaviours, but these latter processes are not elaborated. However, we aim to refine the theory of human functioning and school organisation by considering in more detail the importance of young people's agency, the student peer system and how this interacts with the school institutional system, both to define the ethos of a school and to structure students' health-related actions and outcomes.

The student peer system, like the school institutional system, is guided by a set of social norms, relationships, rituals and symbolic practices, and is also influenced by broader social and structural factors beyond the boundaries of the school. In line with Bernstein¹⁰⁵ and Markham and Aveyard,⁶² we subdivide the institutional features of schools according to their 'instructional' and 'regulatory' functions. However, as our

qualitative synthesis highlights, students not only react to the organisational systems put in place by the institution, but also promote their own versions of these orders. We would suggest that the importance of the student peer system is only partially theorised and that these orders should be reconceptualised to acknowledge more explicitly the student-led as well as institutionally driven nature of the instructional and regulatory orders.

The instructional order of the student peer system is apparent, for example, in the process of learning and adopting symbolic styles of dress and demeanour, and related practices such as violence and substance use. These may be critical to students' safety and social status, particularly in schools that institutionally fail to secure students' engagement or safety. The 'tough' performative rituals consistently reported in the qualitative studies we have reviewed are not indications of 'anarchy' or 'chaos'. Rather they form the symbolic basis of the peer system's own regulatory order. Social control, mediated through rules and norms, seems to be operating among the student body, but on the terms of young people themselves. Take, for example, the rigid rules that students in some schools follow when confronted with a violent incident, such as the 'one-to-one-fight' rule or the practice of linking arms to block adult intrusion. In terms of social capital theory, which our theory synthesis also drew on and which our empirical evidence broadly supports, this also creates an opportunity for students to build bonding social capital¹⁰² and create shared values and obligations which may be critical in school environments that are unable to foster these by other means. In such contexts, students establish their own regulatory order with the aim of facilitating attachment to the peer group and socially regulating these seemingly 'antisocial' risk behaviours. These phenomena can also be understood through the lens of the social development model, which would suggest that participation in ASBs has its own reward structure.

These interpretations resonate with Bourgois'¹⁶⁵ notion of 'street culture' and Dance's¹⁶⁴ concept of 'tough fronts', which conceptualise young people not merely as victims of structural poverty but as agents struggling for meaning and survival, and ultimately reinforcing existing educational and health inequalities. Our argument is that this connection to the peer system's instructional order, as with that to the school's institutional system, enables young people to develop capacity for practical reasoning. In those schools lacking what Markham and Aveyard⁶² term 'institutional authority', this reasoning may provide students with the ability to develop anti-school identities as well as the knowledge and skills required to engage in symbolic practices such as violence and substance use, paradoxically to stay safe and get by. The peer regulatory order also provides a source of affiliation. In schools lacking widespread student connection with the institutional regulatory order, this affiliation may reinforce and regulate violence, substance use and other risk behaviours. Thus, rather than risk stemming from an absence of practical reasoning, affiliation and autonomy, risk actually arises from students developing the autonomy to engage in behaviour that is often regarded as antisocial but which is thoroughly social in its origins.

Finally, it is important to recognise that the way in which the student peer system operates will reflect the way in which the school institutional system operates and the degree to which it is authoritative. The peer system will also in part determine how the formal institution operates. For example, Devine¹⁵¹ illustrated that it is far harder for the institution to be authoritative in a context in which students have little hope of conventional social advancement and are deeply immersed in street culture.

This revised framework therefore builds on Markham and Aveyard's⁶² theory of human functioning and school organisation and the empirical evidence we have reviewed, which suggests that the student population not only reacts to the instructional and regulatory orders established by the school institutional system, but also produces its own set of orders that also shape school organisation and their own risk behaviours. In the qualitative research we have reviewed, the separation of these two systems represents a lack of co-operative functioning of shared norms and understanding between students and the institution, a condition most pervasive in urban contexts of poverty and disadvantage. In this context, students must protect themselves and develop relationships by means of their own intervention. The strategies they adopt, however, create a vicious circle whereby acting 'tough' or 'escaping' the school may lead to aggressive behaviours, poor eating habits, drug and tobacco use and poor sexual and bodily health practices. These

behaviours in turn reinforce the boundaries between students and the authority network, which inhibits the realisation of 'pro-school' reasoning and affiliation for students and ultimately students' capacities for choosing to be healthy.

These suggestions are consistent with the theory of human functioning and school organisation. They are intended to underline the importance of the student peer system in defining a school's ethos in interaction with the formal institutional system, and to describe the ways in which lack of student connection to the institutional instructional and regulatory orders may predispose students' engagement in risky peer groups and behaviours. We acknowledge that the way we have refined Markham and Aveyard's⁶² theory inevitably reflects the range of qualitative research synthesised in this report, and the emphasis on the disconnection between the school institutional and student peer systems is likely to be partly a reflection of the fact that the majority of these studies examined schools in the most deprived urban/inner-city contexts where students and teachers may have the least in common.

Conclusions

We focused on how schools are managed, designed and built and provide learning and teaching, pastoral care and discipline. There is evidence for the potential of school environment interventions addressing these to promote health but the evidence is far from definitive. Five outcome evaluations examined interventions encouraging staff/students to build a stronger sense of community and/or better interpersonal relations in a range of US/UK school settings. These evaluations generally reported benefits, including for measures related to emotional health and aggression. Two evaluations assessed interventions modifying American middle schools' food/physical activity environments and empowering students' involvement in this, reporting benefits for physical activity measures but not for diet. Process evaluations positively reported on the feasibility, fidelity, reach and acceptability of interventions. To develop a fuller picture of the effects of school environment interventions, the results of our own review should be read in conjunction with those of the forthcoming Cochrane review of HPS interventions, which include school environment alongside curriculum and parent/community components.

Outcome and process evaluations were subject to methodological limitations and were not informed by nor aimed at testing any of our review theories. Most of the interventions employed multiple components addressing different aspects of schools' organisation and practice so they do not lend themselves to testing specific hypotheses. However, the evidence from these lends broad support to each of our three primary hypotheses arising from the social development model (regarding the importance for health of participation in school activities), social capital theory (regarding the effects of trusting relationships) and the theory of human functioning and school organisation (regarding the importance of eroding rigid social boundaries between staff and students, and that more student-centred framing of activities will enable better health outcomes).

The multilevel studies provide greater insights regarding our review hypotheses, most notably regarding the theory of human functioning and school organisation, which several studies explicitly aimed to test and provided evidence for.

The meta-ethnography of qualitative studies also supported the theory of human functioning and school organisation, suggesting that a lack of safety at schools, weak student-staff relationships, lack of student participation in decisions and educational disengagement may harm student health.

We have concluded that, although existing interventions suggest the potential for school environment interventions to promote young people's health, the evidence base is currently far from definitive. There is a need for better-conducted RCTs, studies outside the USA, and interventions focused on outcomes other than violence, healthy eating and physical activity. The multilevel studies and qualitative evidence reviewed have suggested potential new foci for intervention studies, such as interventions addressing student

engagement, attainment and attendance, student participation in decisions, and the school physical environment. More trials are also needed to improve the evidence base concerning interventions addressing school community building and interpersonal relationships, particularly in secondary schools and outside the USA. Randomised trials of playground improvements are also required.

Implications

Research implications

1. More empirical research to examine Markham and Aveyard's theory of human functioning and school organisation

The existing literature remains dominated by theories developed to explain crime and ASB, rather than health outcomes. We have attempted to develop a refined theory of human functioning and school organisation, building on Markham and Aveyard's⁶² theory but also drawing on the social development model, theories of social capital and the empirical research reviewed in this report. More research is now needed to examine the theory of human functioning and school organisation and this should go beyond substance use to examine a greater range of health outcomes. This research would include intervention research, multilevel studies of school effects and qualitative research. Future studies need to develop and pilot new measures relating to the constructs outlined in this theory because even where previous empirical studies are explicitly informed by the theory of human functioning and school organisation, they use crude measures. Feminist perspectives could also be integrated to explicitly examine how sex affects school health outcomes and intersects with other factors such as ethnicity and students' SES.¹⁵⁹

2. Intervention studies to address a broader array of school environment factors, for example student attainment, engagement and participation in decisions and the school physical environment

We have concluded that although existing interventions suggest the potential for school environment interventions to promote young people's health the evidence base is currently far from definitive. There is a need for better-conducted RCTs and our review has identified particular gaps, such as the lack of studies outside the USA and the need for interventions focused on outcomes other than violence, healthy eating and physical activity. The other forms of evidence reviewed have suggested several potential new foci for intervention studies.

There is evidence from several well-conducted multilevel studies^{53,64-66} that schools which 'add value' in terms of students' attainment and attendance, and which enable students to participate in decisions, can reduce rates of substance use and violence. Qualitative studies by Astor *et al.*¹⁵³ and Devine¹⁵¹ suggest that schools may be able to foster improved engagement by ensuring that teachers' roles are not limited merely to classroom-based academic teaching but include the pastoral care and discipline of students inside and outside the classroom. Qualitative research by Waldron,¹⁶¹ Plano Clark *et al.*,¹⁵⁵ Lundblad *et al.*¹⁶⁰ and Fletcher *et al.*⁹¹⁻⁹³ suggests that schools may also improve engagement by enabling students to participate in decisions about how schools are run. Students from particular socioeconomic and ethnic groups might be particularly vulnerable to disengagement, also suggested by our qualitative research. Therefore, trials of school environment interventions should examine the effectiveness of increasing student engagement, attainment, attendance and participation in school life.

Evidence from multilevel studies suggests that specific school rules and policies may not always be an effective means of reducing risk behaviours such as substance use in isolation from broader changes to the school environment.¹³¹ Furthermore, qualitative research reviewed here suggests that school rules and policies might sometimes actually exacerbate student antipathy towards staff and disengagement from schools, with the potential to increase investment in behaviours such as smoking, drug use and violence (see, for example, references 91, 155, 160 and 161). Schools might therefore attempt to ensure that school rules and policies take account of student needs, are developed in consultation with students and are

implemented fairly and consistently regardless of factors such as academic and sporting ability, socioeconomic status and ethnic status.

Evidence from qualitative studies by Astor *et al.*¹⁴⁸ and Devine¹⁵¹ as well as a multilevel study by Kumar *et al.*⁶⁸ also suggests that health-risk behaviours can cluster in particular 'unowned spaces' in schools and that the physical environment may influence the quality of social relationships. Schools may therefore be able to address this spatial patterning of risk by reducing such spaces (e.g. ensuring that empty classrooms are locked, that school staff are present in all areas of the school site at break times and lunchtimes). Interventions could involve students and school staff working with architects and designers to ensure that the school physical environment does not inadvertently lead to adverse health problems, evaluated using RCTs. Future intervention studies might examine the potential for interventions addressing these factors.

Some of the evidence we reviewed in relation to RQ2 was of poor quality. More RCTs are needed to improve the evidence base concerning interventions addressing school community building and interpersonal relationships, particularly in secondary schools and outside the USA. Randomised trials of playground improvements are also required.

3. Trials to examine multiple outcomes and health inequalities

When trialling complex school-based health promotion interventions researchers tend to specify just one a priori 'primary outcome'. However, the evidence synthesised in this review suggests that this may not always be the best approach for school environment interventions because multiple health outcomes appear to be potentially affected by such interventions. Furthermore, the meta-ethnography drew attention to how diverse student health outcomes (usually thought to be largely unrelated and the product of different social determinants) may be shaped by common features of the school environment. Furthermore, we would suggest that evaluations of school environment interventions should examine effects on educational as well as health outcomes. This is both because evidence of impacts on attainment may be critical to inform the broader adoption of such interventions and because this would facilitate modelling of the long-term effects on health mediated by increased attainment.¹⁶⁶

We also currently know very little about the potential impact of school environment interventions on health inequalities. Intervention studies should therefore examine effects overall as well as by SES, ethnicity and sex, and should use formal tests for interaction to assess whether or not there are real differences by subgroup.

4. Intervention studies to be more informed by theory

The intervention studies we reviewed generally did not refer to theory. Future outcome and process evaluations of school environment interventions, multilevel studies and qualitative research on school effects on health should employ theory to clarify what processes are being tested and what aspects of the school environment are being examined.

5. Intervention studies to examine cost-effectiveness

There is also a dearth of evidence on the cost-effectiveness of school environment interventions and so future trials should examine this.

6. Intervention studies to include process evaluations

Intervention studies also need to examine processes of planning, delivery and receipt as well as context. Process evaluations should draw both on quantitative data to assess intended processes (e.g. structured observations to assess fidelity, surveys to assess satisfaction) and on qualitative data, which are useful in examining unintended events (e.g. unexpected reasons why intervention delivery is compromised) as well as how context shapes intervention feasibility and acceptability. This was not a common feature of the process evaluations we reviewed.

7. Intervention studies to examine the effects of intervention components in isolation and in combination

We have stressed that the results of our own review, which focuses purely on the effects of the school environment on health, should be read in conjunction with the forthcoming Cochrane review of HPS interventions. HPS interventions include school environment alongside curriculum and parent/community components. Through synergies these might achieve greater effects than school environment-only interventions. Therefore, future trials could provide useful evidence by examining such school environment and curricular/community components both separately and in combination to assess active ingredients. One approach would be to use factorial trials.

8. Multilevel studies to examine a broader range of exposures and outcomes using more rigorous methods

We identified very few multilevel studies which took an approach to adjustment for covariates that enables assessment of school effects on student health. More studies are required to provide more definitive evidence and suggest new avenues for intervention research. Most of the best-conducted multilevel studies were undertaken in the USA and the UK and so there is a need to increase the range of settings examined to explore how school effects vary with local context. Future studies should broaden the health outcomes examined beyond substance use and ensure that they take an appropriate approach to covariates, adjusting for potential confounders but not overadjusting for potential mediators in a mistaken attempt to reduce confounding. Our findings indicate that there is a need for better, more consistent reporting of multilevel studies, for example consistently reporting ICCs. Future multilevel studies should ideally use longitudinal designs, collecting baseline data on young people and their families prior to entry to the school to address the considerable challenge of minimising confounding in such studies. This approach is also required to reduce the extent to which the effects of primary schools confound the apparent effects of secondary schools, a point recognised in the education literature.¹⁶⁷ Longitudinal studies might follow students up for longer periods to explore whether or not any immediate school effects, for example on health behaviours, translate into outcomes related to longer-term lifestyles and chronic disease outcomes. As with school environment interventions, we know little about the impact of school-level factors on health inequalities. Future multilevel studies should also examine subgroup effects and use interaction tests.

9. Multilevel models to examine health effects of different school models

There is also a paucity of studies examining the effects of school type and different models of schooling on health outcomes. In the UK, the USA and elsewhere there is an increasing diversification of school types. In the UK, for example, there is a range of new institutions such as academies and free schools that are funded and accountable to central not local government.¹⁶⁸ No studies we reviewed examined how health outcomes differ between the different school models.

10. Multilevel studies to examine the effects of levels other than school, for example classrooms

Multilevel studies must also, however, avoid reifying the school effect and should assess the contributions of particular classes, year groups and teachers. Our qualitative research suggests that future multilevel studies should examine the effects of peer structures as well as the more formal aspects of schools as institutions.

11. Multilevel studies to be informed by theory

As with intervention studies, future multilevel studies of school effects on health should be informed by theory to enable better interpretation of their results and progressive assessment of the competing theories of how schools influence health. Several otherwise excellent studies we reviewed, such as that by Evans-Whipp *et al.*,¹³¹ do not provide a theoretical rationale for their choice of exposures and this hampers our understanding of exactly what potential causal pathway(s) the studies are aiming to examine. Such problems of empirical research being insufficiently informed by theory have been previously highlighted in public health research.¹⁶⁹

12. Qualitative research to examine broader areas of health, settings and aspects of the school environment

In terms of qualitative research there are a number of clear gaps to be addressed. Most of the qualitative research that we reviewed examined aggression and violence or substance use. There are few conceptually rich studies exploring the potential impacts of the school environment on other public health priorities, such as young people's diet and sexual health, and no qualitative studies were identified that explore the relationship between young people's experience of school and their attitudes and actions relating to physical activity. There is also a lack of studies undertaken outside of the USA and of studies conducted in more socially and economically advantaged urban communities, suburban areas or rural contexts. Existing qualitative research also appears to have neglected how school management and organisation affects student health, and future studies might address this by drawing on data from both students and staff to develop a fuller picture of schools as institutions. Future qualitative studies should also aim to make better use of participatory methods^{170,171} such as mapping spaces and networks and/or peer-led research.

13. Future reviews to examine narrower questions about specific aspects of the school environment and interventions to address these

The majority of studies included in this review were undertaken within the last 10 years, which suggests that this review should be updated every 5 years. Further reviews might differ from our own, however. These might build on our own review by focusing on more homogeneous interventions and school-level influences. These might, for example, divide interventions and school-level determinants into categories such as attainment, classroom management and school design. In some cases, depending on the theories of change underlying interventions or school-level effects, such reviews might focus on a more homogeneous group of health outcomes. For example, future reviews that focus on schools' physical environment might variously be focused on specific outcomes such as physical activity or violence.

Box 1 provides a summary of the implications for research.

Implications for public health

1. Incentives and resources for schools in promoting health

Incentives and resources can be used to ensure that schools are able to promote student health and well-being. Research suggests that some incentive structures such as league tables in the UK and No Child

BOX 1 Summary of implications for research

1. More empirical research to examine Markham and Aveyard's⁶² theory of human functioning and school organisation
2. Intervention studies to address a broader array of school environment factors, e.g. student attainment, engagement, participation in decisions and school physical environment
3. Intervention studies to examine multiple outcomes and health inequalities
4. Intervention studies to be more informed by theory
5. Intervention studies to examine cost-effectiveness
6. Intervention studies to include process evaluations
7. Intervention studies to examine the effects of intervention components in isolation and in combination
8. Multilevel studies to examine a broader range of exposures and outcomes using more rigorous methods
9. Multilevel models to examine health effects of different school models
10. Multilevel studies to examine the effects of levels other than school, e.g. classrooms
11. Multilevel studies to be informed by theory
12. Qualitative research to examine broader areas of health, settings and aspects of the school environment
13. Future reviews to examine narrower questions about specific aspects of the school environment and interventions to address these

Left Behind monitoring systems in the USA can inadvertently increase inequalities in schooling¹⁷² by creating perverse incentives for schools to focus attention on 'key marginal' students who are 'borderline' regarding whether or not they will contribute positively to schools' overall attainment targets, therefore reducing support for those not on course to achieve this.¹⁷³ One means of avoiding such perverse effects might be for the metrics on which schools are judged to focus on measures of the performance of all students (examining means and SDs) rather than measures of the proportion of students achieving a certain targeted threshold, which have recently been introduced in England.

2. School architecture to support promotion of health

Although RCTs should be undertaken to examine the effects of improvements to playgrounds on student health, school building will inevitably occur outside the context of research. The evidence from multilevel studies addressing RQ4 and qualitative research addressing RQ5 is sufficient for us to identify that school buildings which minimise 'unowned' and unsupervised space in school can reduce violence and substance use, and schools which maximise the area for school playgrounds, can maximise physical activity.

Box 2 provides a summary of the implications for public health.

BOX 2 Summary of implications for public health

1. Incentives and resources for schools in promoting health.
2. School architecture to support promotion of health

Acknowledgments

We would like to thank Val Hamilton, Carol Vigurs, Dr Rebecca Langford, Jeff Brunton, Sergio Graziosi, Dr Alison O'Mara-Eves, Zahida Suleman and Juan Daniel Kennedy for their contribution to the project. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the National Institute for Health Research Public Health Research programme or the Department of Health.

This project is funded by a grant from the National Institute for Health Research Public Health Research programme (grant 09/3002/08).

Contribution of authors

Chris Bonell planned the project, led on synthesising theory, outcome evaluations, process evaluations and multilevel studies and led on report drafting.

Farah Jamal conducted the review of qualitative research and contributed towards report drafting.

Angela Harden contributed to planning the project and led on synthesising qualitative research.

Helene Wells managed the review day-to-day and contributed to synthesising theory, outcome evaluations, process evaluations and multilevel studies and report drafting.

Will Parry reviewed the multilevel studies and contributed to report drafting.

Adam Fletcher contributed to project planning, synthesis of qualitative research and report drafting.

Mark Petticrew contributed to project planning and commented on a report draft.

James Thomas contributed to project planning, advised on ER4 and multilevel research and commented on two report drafts.

Margaret Whitehead contributed to project planning and commented on a report draft.

Rona Campbell contributed to project planning and commented on a report draft.

Simon Murphy contributed to project planning and commented on a report draft.

Laurence Moore contributed to project planning and advised on multilevel studies.

References

1. Donaldson L. *Tackling the health of the teenage nation: Chief Medical Officer's annual report 2007*. London: Department of Health; 2008.
2. UNICEF. *An overview of child wellbeing in rich countries: a comprehensive assessment of the lives and wellbeing of children and adolescents in the economically advanced nations*. Florence: Innocenti Research Centre; 2007.
3. Jackson C, Sweeting H, Haw S. Clustering of substance use and sexual risk behaviour in adolescence: analysis of two cohort studies. *BMJ Open* 2012;**2**:e000661. <http://dx.doi.org/10.1136/bmjopen-2011-000661>
4. Resnick MD, Bearman PS, Blum RW, Bauman KE, Harris KM, Jones J, et al. Protecting adolescents from harm: findings from the national longitudinal study on adolescent harm. *J Am Med Assoc* 1997;**278**:823–32. <http://dx.doi.org/10.1001/jama.278.10.823>
5. DiCenso A, Guyatt G, Willan A, Griffith L. Interventions to reduce unintended pregnancies among adolescents: systematic review of randomised controlled trials. *BMJ* 2002;**324**:1426–34.
6. Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P. School-based prevention for illicit drugs use. *Cochrane Database Syst Rev* 2005;**2**:CD003020.
7. Foxcroft DR, Ireland D, Lowe G, Breen R. Primary prevention for alcohol misuse in young people. *Cochrane Database Syst Rev* 2002;**3**:CD003024.
8. Harden A, Oakley A, Oliver S. Peer-delivered health promotion for young people: a systematic review of different study designs. *Health Educ J* 2001;**60**:339–53.
9. Oliver S, Kavanagh J, Caird J, Lorenc T, Oliver K, Harden A, et al. *Health promotion, inequalities and young people's health: a systematic review of research*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; 2008.
10. Thomas R, Perera R. School-based programmes for preventing smoking. *Cochrane Database Syst Rev* 2006;**3**:CD001293.
11. Wells J, Barlow J, Stewart-Brown S. A systematic review of universal approaches to mental health promotion in schools. *Health Educ* 2003;**103**:197–220.
12. Bronfenbrenner U. *The ecology of human development: experiments by nature and design*. Cambridge, MA: Harvard University Press; 1979.
13. Brooks-Gunn J, Duncan G, Aber L, editors. *Neighborhood poverty*. Vol. 1. New York, NY: Russell Sage Foundation; 1997.
14. Rutter M, Maughan B, Mortimore P, Ouston J. *Fifteen thousand hours*. London: Open Books; 1979.
15. Coleman J, Campbell E, Hobson C, McPartland J, Mood A, Weinfield F, et al. *Equality of educational opportunity*. Washington, DC: US Government Printing Office; 1966.
16. MacBeath J, Mortimore P. *Improving school effectiveness*. Buckingham: Open University Press; 2001.
17. Arnot M, Gray J, James M, Rudick J, Duveen G. *Recent research on gender and educational performance*. London: The Stationery Office; 1998.
18. Gaine C, George R. *Gender, race and class in schooling: a new introduction*. London: Falmer Press; 1999.

19. Gripps C, Murphy P. *A fair test? Assessment, achievement and equity*. Buckingham: Open University Press; 1994.
20. Scheerens J. *Improving school effectiveness. Fundamentals of education planning*. Paris: UNESCO; 2000.
21. MacBeath J, Thomson B, Arrowsmith J, Forbes D. *Using ethos indicators in secondary school self-evaluation: taking account of pupils, parents and teachers*. Edinburgh: Scottish Office Education Department, HM Inspectors of Schools; 1992.
22. Ulrich C. A place of their own: children and the physical environment. *Hum Ecol* 2004;**32**:11–14.
23. Murray M, Kiryluk S, Swan AV. School characteristics and adolescent smoking: results from the MRC/Derbyshire Smoking Study. *J Epidemiol Community Health* 1984;**38**:167–72.
24. Cuijpers P, Jonkers R, de Weerd I, de Jong A. The effects of drug abuse prevention at school: the 'Healthy School and Drugs' project. *Addiction* 2002;**97**:67–73.
25. Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, et al. *Young people and healthy eating: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit; 2001.
26. Rees R, Harden A, Shepherd J, Brunton G, Oliver S, Oakley A. *Young people and physical activity: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit; 2001.
27. Bond L, Carlin JB, Thomas L, Rubin K, Patton G. Does bullying cause emotional problems? A retrospective study of young teenagers. *BMJ* 2001;**323**:480–4.
28. Catalano RF, Haggerty KP, Oesterle S, Fleming CB, Hawkins JD. The importance of bonding to school for healthy development: findings from the Social Development Research Group. *J Sch Health* 2004;**74**:252–61.
29. Dahlgren G, Whitehead M. *Policies and strategies to promote social equity in health*. Stockholm: Institute for Futures Studies; 1991.
30. Marmot MG. Creating healthier societies. *Bull World Health Org* 2004;**82**:320–1.
31. World Health Organization. *Ottawa Charter for Health Promotion*. Copenhagen: WHO Regional Office for Europe; 1986.
32. Health Education Board for Scotland HPW, Health Education Authority (England), Health Promotion Agency for Northern Ireland. The European Network of Health Promoting Schools: introduction – the UK project. *Health Educ J* 1996;**55**:447–9.
33. Denman S, Moon A, Parsons C, Stears D. *The health promoting school. Policy, research and practice*. London: Routledge/Farmer; 2002.
34. Lister-Sharp DJ, Chapman S, Stewart-Brown S, Sowden A. Health promoting schools and health promotion in schools: two systematic reviews. *Health Technol Assess* 1999;**3**(22).
35. Nutbeam D. The health promoting school: closing the gap between theory and practice. *Health Promot Int* 1992;**7**:151–3.
36. Parsons C, Stears D, Thomas C, Holland J. The health promoting school in Europe: conceptualising and evaluating the challenge. *Health Educ J* 1996;**55**:311–21.
37. World Health Organization. *Promoting health through schools*. Geneva: WHO; 1997.
38. Young I, Williams T. *The healthy school*. Edinburgh: Scottish Health Education Group; 1989.

39. Lynagh M, Schofield J, Sanson-Fisher R. School health promotion programmes over the past decade: a review of the smoking, alcohol and solar protection literature. *Health Promot Int* 1997;**12**:43–60. <http://dx.doi.org/10.1093/heapro/12.1.43>
40. Leger LS. Australian teachers' understandings of the health promoting school concept and the implications for the development of school health. *Health Promot Int* 1998;**13**:223–35.
41. Wilson D, Gottfredson DC, Najaka S. School-based prevention of problem behaviors: a meta-analysis. *J Quant Criminol* 2001;**17**:241–71.
42. Fletcher A, Bonell C, Hargreaves J. School effects on young people's drug use: a systematic review of intervention and observational studies. *J Adolesc Health* 2008;**42**:209–20. <http://dx.doi.org/10.1016/j.jadohealth.2007.09.020>
43. Bond L, Thomas L, Coffey C, Glover S. Long-term impact of the Gatehouse Project on cannabis use of 16-year-olds in Australia. *J School Health* 2004;**74**:23–30. <http://dx.doi.org/10.1111/j.1746-1561.2004.tb06597.x>
44. Flay BR, Graumlich S, Segawa E, Burns JL, Holliday MY. Effects of 2 prevention programs on high-risk behaviors among African American youth: a randomized trial. *Arch Pediatr Adolesc Med* 2004;**158**:377–84. <http://dx.doi.org/10.1001/archpedi.158.4.377>
45. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Integrating process evaluation in the design of randomised controlled trials of complex interventions: the example of the RIPPLE study. *BMJ* 2006;**332**:413–16.
46. Patton G, Bond L, Carlin JB, Thomas L, Butler H, Glover S, *et al*. Promoting social inclusion in schools: group-randomized trial of effects on student health risk behaviour and well-being. *Am J Public Health* 2006;**96**:1582–7. <http://dx.doi.org/10.2105/AJPH.2004.047399>
47. Evans-Whipp T, Beyers JM, Lloyd S, Lafazia AN, Toumbourou JW, Arthur MW, *et al*. A review of school drug policies and their impact on youth substance use. *Health Promot Int* 2004;**19**:227–34. <http://dx.doi.org/10.1093/heapro/dah210>
48. Aveyard A, Markham WA, Chenge KK. A methodological and substantive review of the evidence that schools cause pupils to smoke. *Soc Sci Med* 2004;**58**:2252–65. <http://dx.doi.org/10.1016/j.socscimed.2003.08.012>
49. Sellström E, Bremberg S. Is there a 'school effect' on pupil outcomes? A review of multilevel studies. *J Epidemiol Community Health* 2006;**60**:149–55. <http://dx.doi.org/10.1136/jech.2005.036707>
50. Vreeman RC, Carroll AE. A systematic review of school-based interventions to prevent bullying. *Arch Pediatr Adolesc Med* 2007;**161**:78–88. <http://dx.doi.org/10.1001/archpedi.161.1.78>
51. Langford R, Campbell R, Magnus D, Bonell CP, Murphy SM, Waters E, *et al*. The WHO Health Promoting School Framework for improving the health and well-being of students and staff. *Cochrane Database Syst Rev* 2011;**1**:CD008958. <http://dx.doi.org/10.1002/14651858.CD008958>
52. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4.0: software for research synthesis*. EPPI-Centre software. London: Social Science Research Unit, Institute of Education, University of London; 2010.
53. Tobler AL, Komro KA, Dabroski A, Aveyard P, Markham WA. Preventing the link between SES and high-risk behaviors: 'value-added' education, drug use and delinquency in high-risk, urban schools. *Prev Sci* 2011;**12**:211–21. <http://dx.doi.org/10.1007/s11121-011-0206-9>
54. Battistich V, Solomon D, Delucchi KL. Effects of a program to enhance prosocial development on adjustment. Annual Meeting of the American Psychological Association, Boston, MA, 10–14 August 1990.

55. Battistich V, Solomon D, Watson M, Solomon J, Schaps E. Effects of an elementary school program to enhance prosocial behavior on children's cognitive-social problem solving skills and strategies. *J Appl Dev Psychol* 1989;**10**:147–69. [http://dx.doi.org/10.1016/0193-3973\(89\)90002-6](http://dx.doi.org/10.1016/0193-3973(89)90002-6)
56. Flay BR. Understanding environmental, situational and intrapersonal risk and protective factors for youth tobacco use: the theory of triadic influence. *Nicotine Tob Res* 1999;**1**(Suppl. 2):S111–14. <http://dx.doi.org/10.1080/14622299050011911>
57. Lohrmann DK. A complementary ecological model of the coordinated school health program. *Public Health Rep* 2008;**123**:695–703. <http://dx.doi.org/10.1111/j.1746-1561.2009.00460.x>
58. Solomon D, Watson MS, Delucchi KL, Schaps E, Battistich V. Enhancing children's prosocial behavior in the classroom. *Am Educ Res J* 1988;**25**:527–54. <http://dx.doi.org/10.2307/1163128>
59. Bonell C, Sorhaindo A, Strange VJ, Wiggins M, Allen E, Fletcher A, et al. A pilot whole-school intervention to improve school ethos and reduce substance use. *Health Educ* 2010;**110**:252–72. <http://dx.doi.org/10.1108/96542831080001390>
60. Bonell C, Sorhaindo A, Allen ES, Strange VJ, Wiggins M, Fletcher A, et al. Pilot multimethod trial of a school-ethos intervention to reduce substance use: building hypotheses about upstream pathways to prevention. *J Adolesc Health* 2010;**47**:555–63. <http://dx.doi.org/10.1016/j.jadohealth.2010.04.011>
61. Dobbins M, De Corby K, Robeson P, Husson H, Tirilis D. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6–18. *Cochrane Database Syst Rev* 2009;**1**:CD007651. <http://dx.doi.org/10.1002/14651858.CD007651>
62. Markham WA, Aveyard P. A new theory of health promoting schools based on human functioning, school organisation and pedagogic practice. *Soc Sci Med* 2003;**56**:1209–20. [http://dx.doi.org/10.1016/S0277-9536\(02\)00120-X](http://dx.doi.org/10.1016/S0277-9536(02)00120-X)
63. Dzewaltowski DA, Estabrooks PA, Welk G, Hill J, Milliken G, Karteroliotis K, et al. Healthy youth places: a randomized controlled trial to determine the effectiveness of facilitating adult and youth leaders to promote physical activity and fruit and vegetable consumption in middle schools. *Health Educ Behav* 2009;**36**:583–600. <http://dx.doi.org/10.1177/1090198108314619>
64. Aveyard P, Markham WA, Lancashire E, Bullock A, MacArthur C, Cheng KK, et al. The influence of school culture on smoking among pupils. *Soc Sci Med* 2004;**58**:1767–80. [http://dx.doi.org/10.1016/S0277-9536\(03\)00396-4](http://dx.doi.org/10.1016/S0277-9536(03)00396-4)
65. Markham WA, Aveyard P, Bisset SL, Lancashire ER, Bridle C, Deakin S. Value-added education and smoking uptake in schools: a cohort study. *Addiction* 2008;**103**:155–61. <http://dx.doi.org/10.1111/j.1360-0443.2007.02020.x>
66. Bisset S, Markham WA, Aveyard P. School culture as an influencing factor on youth substance use. *J Epidemiol Community Health* 2007;**61**:485–90. <http://dx.doi.org/10.1136/jech.2006.048157>
67. Henry KL, Stanley LR, Edwards RW, Harkabus LC, Chapin LA. Individual and contextual effects of school adjustment on adolescent alcohol use. *Prev Sci* 2009;**10**:236–47. <http://dx.doi.org/10.1007/s11121-009-0124-2>
68. Kumar R, O'Malley PM, Johnston LD. Association between physical environment of secondary schools and student problem behavior: a national study, 2000–2003. *Environ Behav* 2008;**40**:455–86. <http://dx.doi.org/10.1177/0013916506293987>
69. Monshouwer K, Van Dorsselaer S, Van Os J, Drukker M, De Graaf R, Ter Bogt T, et al. Ethnic composition of schools affects episodic heavy drinking only in ethnic-minority students. *Addiction* 2007;**102**:722–9. <http://dx.doi.org/10.1111/j.1360-0443.2007.01760.x>

70. Akiba M. What predicts fear of school violence among us adolescents? *Teach Coll Rec* 2010;**112**:68–102.
71. Johnson RA, Hoffmann JP. Adolescent cigarette smoking in US racial/ethnic subgroups: findings from the National Education Longitudinal Study. *J Health Soc Behav* 2000;**41**:392–407.
72. Collins D, Pan Z, Johnson K, Courser M, Shamblen S. Individual and contextual predictors of inhalant use among 8th graders: a multilevel analysis. *J Drug Educ* 2008;**38**:193–210. <http://dx.doi.org/10.2190/DE.38.3.a>
73. Rountree PW, Clayton RR. A contextual model of adolescent alcohol use across the rural–urban continuum. *Subst Use Misuse* 1999;**34**:495–519. <http://dx.doi.org/10.3109/10826089909037228>
74. Wiium N, Wold B. Family and school influences on adolescent smoking behaviour. *Health Educ* 2006;**106**:465–79. <http://dx.doi.org/10.1108/09654280610711415>
75. Wiium N, Wold B. An ecological system approach to adolescent smoking behavior. *J Youth Adolesc* 2009;**38**:1351–63. <http://dx.doi.org/10.1007/s10964-008-9349-9>
76. Gastic B. *Student attitudes about school safety and their implications for safe school reform*. PhD thesis. Stanford, CA: Stanford University; 2005.
77. Haley SJ. *School context and the association of risk and protective factors among same-sex attracted youth*. Dissertation. Waltham, MA: Brandeis University; 2007.
78. Khoury-Kassabri M, Astor RA, Benbenishty R. Middle Eastern adolescents' perpetration of school violence against peers and teachers: a cross-cultural and ecological analysis. *J Interpers Violence* 2009;**24**:159–82. <http://dx.doi.org/10.1177/0886260508315777>
79. Kairouz S, Adlaf EM. Schools, students and heavy drinking: a multilevel analysis. *Addict Res Theory* 2003;**11**:427–39. <http://dx.doi.org/10.1080/1606635021000058485>
80. Lothian SF. *Understanding student smoking behavior: an ecological approach*. PhD thesis. London, ON: University of Western Ontario; 2003.
81. Kim S. *Prevention of adolescent substance use: an investigation of cross-level interaction effects of self, peer, family and school level risk and protective factors using multilevel modeling*. PhD thesis. Los Angeles, CA: University of Southern California; 1999.
82. Way SM. *For their own good? The effects of school discipline and disorder on student behavior and academic achievement*. Dissertation. Tucson, AZ: University of Arizona; 2003.
83. Ishibashi SLW. *Lessons learned from tobacco control: a multilevel analysis of school characteristics and adolescent physical activity*. PhD thesis. Waterloo, ON: University of Waterloo; 2009.
84. Murnaghan DA, Sihvonen M, Leatherdale ST, Kekki P. The relationship between school-based smoking policies and prevention programs on smoking behavior among grade 12 students in Prince Edward Island: a multilevel analysis. *Prev Med* 2007;**44**:317–22. <http://dx.doi.org/10.1016/j.ypmed.2007.01.003>
85. Comeau C. *Impact of school context on the relations between deviant peer affiliations and problem behaviors during the middle school years: an exploratory analysis using hierarchical linear modeling*. PhD thesis. Eugene, OR: University of Oregon; 2005.
86. Gladden RM. *Building school communities and making schools safer*. PhD thesis. New York, NY: City University of New York; 2005.
87. Haug E, Torsheim T, Samdal O. Local school policies increase physical activity in Norwegian secondary schools. *Health Promot Int* 2010;**25**:63–72. <http://dx.doi.org/10.1093/heapro/dap040>

88. Kim YS. *The effects of individual, neighborhood, and school characteristics on delinquency: a multi-level cross-nested analysis*. PhD thesis. University Park, PA: Pennsylvania State University; 2001.
89. Bradshaw CP, Sawyer AL, O'Brennan LM. A social disorganization perspective on bullying-related attitudes and behaviors: the influence of school context. *Am J Community Psychol* 2009;**43**: 204–20. <http://dx.doi.org/10.1007/s10464-009-9240-1>
90. Leatherdale ST, Manske S, Faulkner G, Arbour K, Bredin C. A multi-level examination of school programs, policies and resources associated with physical activity among elementary school youth in the play-on study. *Int J Behav Nutr Phys Act* 2010;**7**:1–13. <http://dx.doi.org/10.1186/1479-5868-7-6>
91. Fletcher A, Bonell C, Rhodes T. New counter-school cultures: female students' drug use at a high-achieving secondary school. *Br J Sociol Educ* 2009;**30**:549–62. <http://dx.doi.org/10.1080/01425690903101049>
92. Fletcher A, Bonell C, Sorhaindo A, Rhodes T. Cannabis use and 'safe' identities in an inner-city school risk environment. *Int J Drug Policy* 2009;**20**:244–50. <http://dx.doi.org/10.1016/j.drugpo.2008.08.006>
93. Fletcher A, Bonell C, Sorhaindo A, Strange V. How might schools influence young people's drug use? Development of theory from qualitative case-study research. *J Adolesc Health* 2009; **45**:126–32. <http://dx.doi.org/10.1016/j.jadohealth.2008.12.021>
94. Brotman JS. *Urban high school students' talk about HIV/AIDS decision-making: learning, identities, and the influence of school*. PhD thesis. New York, NY: Columbia University; 2009.
95. Gordon J, Turner K. School staff as exemplars – where is the potential? *Health Educ* 2001;**101**: 283–91. <http://dx.doi.org/10.1108/EUM0000000005991>
96. Marsiglia FF, Miles BW, Dustman P, Sills S. Ties that protect: an ecological perspective on latino/a urban pre-adolescent drug use. *J Ethn Cult Divers Soc Work* 2002;**11**:191–220. http://dx.doi.org/10.1300/J051v11n03_03
97. Grunseit AC, Weatherburn D, Donnelly N. Correlates of physical violence at school: a multilevel analysis of Australian high school students. *Aust J Soc Issues* 2008;**43**:527–45.
98. Hawkins JD, Weis JG. The social development model: an integrated approach to delinquency prevention. *J Prim Prev* 1985;**6**:73–97. <http://dx.doi.org/10.1007/BF01325432>
99. Portes A. Social capital: its origins and applications in modern sociology. *Ann Rev Sociol* 1998; **24**:1–24. <http://dx.doi.org/10.1146/annurev.soc.24.1.1>
100. Putnam RD. *Bowling alone: the collapse and revival of American community*. New York, NY: Simon & Schuster; 2000.
101. Bourdieu P, Wacquant L. *An invitation to reflexive sociology*. Chicago, IL: University of Chicago Press; 1992.
102. Coleman JS. Social capital in the creation of human capital. *Am J Sociol* 1989;**94**(Suppl.):S95–120. <http://dx.doi.org/10.1086/228943>
103. Granovetter M. The strength of weak ties. *Am J Sociol* 1973;**78**:1360–80. <http://dx.doi.org/10.1086/225469>
104. Nussbaum MC. Aristotelian social democracy. In Douglas RB, Mara GM, Richardson H, editors. *Liberalism and the good*. London: Routledge; 1990. pp. 203–51.
105. Bernstein B. *Class, codes and control, vol. 3: towards a theory of educational transmission*. London: Routledge; 1975.

106. Bernstein B. *Pedagogy, symbolic control and identity*. London: Taylor & Francis; 1996.
107. Gottfredson GD, Hirschi T. *A general theory of crime*. Stanford CA: Stanford University Press; 1990.
108. Elliott DS, Ageton SS, Canter RJ. An integrated theoretical perspective on delinquent behavior. *J Res Crime Delinq* 1979;**16**:3–27. <http://dx.doi.org/10.1177/002242787901600102>
109. Jessor R, Jessor S. *Problem behavior and psychosocial development: a longitudinal study of youth*. New York, NY: Academic Press; 1977.
110. Merton RK. Social structure and anomie. *Am Soc Rev* 1938;**3**:672–82. <http://dx.doi.org/10.2307/2084686>
111. Gibbs J. *Crime, punishment and deterrence*. New York, NY: Elsevier; 1975.
112. Ajzen I, Fishbein M. *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall; 1980.
113. Ajzen I. From intentions to actions: a theory of planned behavior. In Kuhl J, Beckman J, editors. *Action-control: from cognition to behavior*. Heidelberg: Springer; 1985. pp. 11–39.
114. Durkheim E. *The elementary forms of religious life*, translated by KE Fields. New York, NY: The Free Press; 1995 [1912].
115. Shepherd J, Kavanagh J, Picot K, Cooper A, Harden E, Barnett-Page J, *et al*. The effectiveness and cost-effectiveness of behavioural interventions for the prevention of sexually transmitted infections in young people aged 13 to 19: a systematic review and economic evaluation. *Health Technol Assess* 2010;**14**(7).
116. Valentine JC, Pigott TD, Rothstein HR. How many studies do you need? A primer on statistical power for meta-analysis. *J Educ Behav Stat* 2010;**35**:215–47. <http://dx.doi.org/10.3102/1076998609346961>
117. Stratton G. Promoting children's physical activity in primary school: an intervention study using playground markings. *Ergonomics* 2000;**43**:1538–46. <http://dx.doi.org/10.1080/001401300750003961>
118. Stratton G, Mullan E. The effect of multicolor playground markings on children's physical activity level during recess. *Prev Med* 2005;**41**:828–33. <http://dx.doi.org/10.1016/j.ypmed.2005.07.009>
119. Flannery DJ, Vazsonyi AT, Liau AK, Guo S, Powell KE, Atha H, *et al*. Initial behavior outcomes for the peacebuilders universal school-based violence prevention program. *Dev Psychol* 2003;**39**:292–308. <http://dx.doi.org/10.1037//0012-1649.39.2.292>
120. Fonagy P, Twemlow SW, Vernberg EM, Nelson JM, Dill EJ, Little TD, *et al*. A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *J Child Psychol Psychiatry* 2009;**50**:607–16. <http://dx.doi.org/10.1111/j.1469-7610.2008.02025.x>
121. Krug EG, Brener ND, Dahlberg LL, Ryan GW, Powell KE. The impact of an elementary school-based violence prevention program on visits to the school nurse. *Am J Prev Med* 1997;**13**:459–63.
122. Sallis JF, McKenzie TL, Conway TL, Elder JP, Prochaska JJ, Brown M, *et al*. Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. *Am J Prev Med* 2003;**24**:209–17.
123. Battistich V. Effects of a school-based program to enhance prosocial development on children's peer relations and social adjustment. *J Res Character Educ* 2003;**1**:1–7.

124. Ridgers ND, Stratton G, Fairclough SJ, Twisk JWR. Long-term effects of a playground marking and physical structures on children's recess physical activity levels. *Prev Med* 2007;**44**:393–7. <http://dx.doi.org/10.1016/j.ypmed.2007.01.009>
125. Ridgers ND, Stratton GS, Fairclough SJ, Twisk JWR. Children's physical activity levels during school recess: a quasi-experimental intervention study. *Int J Behav Nutr Phys Act* 2007;**4**:1–9. <http://dx.doi.org/10.1186/1479-5868-4-19>
126. Ridgers ND, Fairclough SJ, Stratton G. Twelve-month effects of a playground intervention on children's morning and lunchtime recess physical activity levels. *J Phys Act Health* 2010;**7**:167–75.
127. Twemlow SW, Fonagy P, Sacco FC, Gies ML, Evans R, Ewbank R. Creating a peaceful school learning environment: a controlled study of an elementary school intervention to reduce violence. *Am J Psychiatry* 2001;**158**:808–10. <http://dx.doi.org/10.1176/appi.ajp.158.5.808>
128. Bonell C, Oakley A, Hargreaves J, Strange V, Rees R. Trials of health interventions and empirical assessment of generalizability: suggested framework and systematic review. *BMJ* 2006;**333**:346–9.
129. Fagen MC, Flay BR. Effects of a school-based program to enhance prosocial development on children's peer relations and social adjustment. *Health Educ Behav* 2009;**36**:9–23.
130. West P, Sweeting H, Leyland A. School effects on pupils health behaviours: evidence in support of the health promoting school. *Res Papers Educ* 2004;**19**:261–91.
131. Evans-Whipp TJ, Bond L, Ukoumunne OC, Toumbourou JW, Catalan RF. The impact of school tobacco policies on student smoking in Washington State, United States and Victoria, Australia. *Int J Environ Res Public Health* 2010;**7**:698–710. <http://dx.doi.org/10.3390/ijerph7030698>
132. Piontek D, Buehler A, Donath C, Floeter S, Rudolph U, Metz K, et al. School context variables and students' smoking. *Eur Addict Res* 2008;**14**:53–60. <http://dx.doi.org/10.1159/000110411>
133. Cradock AL, Melly SJ, Allen JG, Morris JS, Gortmaker SL. Characteristics of school campuses and physical activity among youth. *Am J Prev Med* 2007;**33**:106–13. <http://dx.doi.org/10.1016/j.amepre.2007.04.009>
134. Adams ML, Jason LA, Pokorny S, Hunt Y. The relationship between school policies and youth tobacco use. *J School Health* 2009;**79**:17–23. <http://dx.doi.org/10.1111/j.1746-1561.2008.00369.x>
135. Ma X. Health outcomes of elementary school students in New Brunswick. The education perspective. *Eval Rev* 2000;**24**:435–56. <http://dx.doi.org/10.1177/0193841X0002400501>
136. Moore L, Roberts C, Tudor-Smith C. School smoking policies and smoking prevalence among adolescents: a multi-level analysis of cross-sectional data from Wales. *Tob Control* 2001;**10**:117–23.
137. West P, Sweeting H, Young R. Transition matters: pupils' experiences of the primary–secondary school transition in the west of Scotland and consequences for well-being and attainment. *Res Papers Educ* 2010;**25**:21–50. <http://dx.doi.org/10.1080/02671520802308677>
138. Attar-Schwartz S, Khoury-Kassabri M. Indirect versus verbal forms of victimization at school: the contribution of student, family, and school variables. *Soc Work Res* 2008;**32**:159–70. <http://dx.doi.org/10.1093/swr/32.3.159>
139. van der Horst K, Oenema A, Velde T, Brug J. Gender, ethnic and school type differences in overweight and energy balance-related behaviours among Dutch adolescents. *Int J Pediatr Obes* 2009;**4**:371–80.
140. Maes L, Lievens J. Adolescent smoking behaviour and the school: a multi-level analysis. In Tudor-Smith C, editor. *Tackling tobacco*. Cardiff: Health Promotion Wales; 1999. pp. 113–28.

141. Pinilla J, González B, Barber P, Santana Y. Smoking in young adolescents: an approach with multilevel discrete choice models. *J Epidemiol Community Health* 2002;**56**:227–32. <http://dx.doi.org/10.1136/jech.56.3.227>
142. Malikaew P, Watt RG, Sheiham A. Associations between school environments and childhood traumatic dental injuries. *Oral Health Prev Dent* 2003;**1**:255–66.
143. Henderson M, Ecob R, Wight D, Abraham C. What explains between-school differences in rates of smoking? *BMC Public Health* 2008;**8**:1–16. <http://dx.doi.org/10.1186/1471-2458-8-218>
144. Noblit G, Hare R. *Meta-ethnography: synthesizing qualitative studies*. London: Sage; 1988.
145. Britten N, Campbell R, Pope C, Donovan J, Morgan M. Using meta-ethnography to synthesise qualitative research: a worked example. *J Health Serv Res Policy* 2002;**7**:4209–15. <http://dx.doi.org/10.1258/135581902320432732>
146. Garside R, Britten N, Stein K. The experience of heavy menstrual bleeding: a systematic review and meta-ethnography of qualitative studies. *J Adv Nurs* 2008;**63**:550–62.
147. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, *et al*. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Soc Sci Med* 2003;**56**:671–84. [http://dx.doi.org/10.1016/S0277-9536\(02\)00064-3](http://dx.doi.org/10.1016/S0277-9536(02)00064-3)
148. Astor RA, Meyer HA, Behre WJ. Unowned places and times: maps and interviews about violence in high schools. *Am Educ Res J* 1999;**36**:3–42. <http://dx.doi.org/10.2307/1163504>
149. Burnett C. School violence in an impoverished South African community. *Child Abuse Negl* 1998;**22**:789–95. [http://dx.doi.org/10.1016/S0145-2134\(98\)00058-1](http://dx.doi.org/10.1016/S0145-2134(98)00058-1)
150. Brunson RK, Miller JM. Schools, neighborhoods, and adolescent conflicts: a situational examination of reciprocal dynamics. *Justice Q* 2009;**26**:183–210. <http://dx.doi.org/10.1080/07418820802245060>
151. Devine J. Can metal detectors replace the panopticon? *Cult Anthropol* 1995;**10**:171–95. <http://dx.doi.org/10.1525/can.1995.10.2.02a00020>
152. Wilson-Simmons R, Dash K, Tehranifar P, O'Donnell L, Stueve A. What can student bystanders do to prevent school violence? Perceptions of students and school staff. *J Sch Violence* 2006;**5**:43–62.
153. Astor RA, Meyer HA, Pitner RO. Elementary and middle school students' perceptions of violence-prone school subcontexts. *Elem Sch J* 2001;**101**:511–28. <http://dx.doi.org/10.1086/499685>
154. Fletcher A, Bonell C. 'Escaping from prison': Female students' substance use and the school environment. *J Epidemiol Community Health* 2008;**62**(Suppl. 1):A10.
155. Plano Clark VL, Miller DL, Creswell JW, McVea K, McEntarffer R, Harter LM, *et al*. In conversation: high school students talk to students about tobacco use and prevention strategies. *Qual Health Res* 2002;**12**:1264–83. <http://dx.doi.org/10.1177/1049732302238249>
156. Wills W, Backett-Milburn K, Gregory S, Lawton J. The influence of the secondary school setting on the food practices of young teenagers from disadvantaged backgrounds in Scotland. *Health Educ Res* 2005;**20**:458–65. <http://dx.doi.org/10.1093/her/cyg132>
157. Haselswerdt MV, Lenhardt AMC. Reframing school violence: listening to voices of students. *Educ Forum* 2003;**67**:326–36. <http://dx.doi.org/10.1080/00131720308984581>
158. Hosie ACS. 'I hated everything about school': an examination of the relationship between dislike of school, teenage pregnancy and educational disengagement. *Soc Policy Soc* 2007;**6**:333–47. <http://dx.doi.org/10.1017/S1474746407003661>
159. Keddie A. 'Some of those girls can be real drama queens': issues of gender, sexual harassment and schooling. *Sex Educ* 2009;**9**:1–16. <http://dx.doi.org/10.1080/14681810802639863>

160. Lundblad B, Hellstrom AL, Berg M. Children's experiences of attitudes and rules for going to the toilet in school. *Scand J Caring Sci* 2010;**24**:219–23. <http://dx.doi.org/10.1111/j.1471-6712.2009.00707.x>
161. Waldron LM. The messy nature of discipline and zero tolerance policies: negotiating safe school environments among inconsistencies, structural constraints and the complex lives of youth. *Sociol Stud Child Youth* 2005;**11**:81–114. [http://dx.doi.org/10.1016/S1537-4661\(05\)11004-6](http://dx.doi.org/10.1016/S1537-4661(05)11004-6)
162. Cousins LH. Toward a sociocultural context for understanding violence and disruption in black urban schools and communities. *J Sociol Soc Welf* 1997;**24**:41–63. [http://dx.doi.org/10.1016/S1537-4661\(05\)11004-6](http://dx.doi.org/10.1016/S1537-4661(05)11004-6)
163. Pike J, Colquhoun D. The relationship between policy and place: the role of school meals in addressing health inequalities. *Health Sociol Rev* 2009;**18**:50–60. <http://dx.doi.org/10.5172/hesr.18.1.50>
164. Dance J. *Tough fronts: the impact of street culture on schooling*. New York, NY: Routledge/Farmer; 2002.
165. Bourgois P. *In search of respect: selling crack in El Barrio*. New York, NY: Cambridge University Press; 1995.
166. Feinstein L. *Quantitative estimates of the social benefits of learning*. London: Centre for Research on the Wider Benefits of Learning, Institute of Education; 2002.
167. Goldstein H, Sammons P. The influence of secondary and junior schools on sixteen year examination performance: a cross-classified multi-level analysis. *School effectiveness and school improvement*. *Sch Eff Sch Improv* 1997;**8**:219–30.
168. Ball SJ. *The education debate: policy and politics in the 21st century*. Bristol: Policy Press; 2008.
169. Frohlich KL, Mykhalovskiy E, Miller F, Daniel MI. Advancing the population health agenda: encouraging the integration of social theory into population health research and practice. *Can J Public Health* 2004;**95**:392–5.
170. Anderson P, Morrow V. *Ethics, social research and consulting children and young people*. Barking, Essex: Barnardo's; 2004.
171. Punch S. Interviewing strategies with young people: the 'secret box', stimulus material and task based activities. *Child Soc* 2002;**16**:45–56. <http://dx.doi.org/10.1002/chi.685>
172. Ball S. *Class Strategies and the Education Market: the Middle Classes and Social Advantage*. London: Routledge Falmer; 2003.
173. Bonell C, Fletcher A, Sorhaindo A, Welles H, McKee M. How market-oriented education policies might influence young people's health: development of a logic model from qualitative case studies in English secondary schools [published online ahead of print 28 June 2011]. *J Epidemiol Community Health* 2011. <http://dx.doi.org/10.1136/jech.2011.137539>
174. Durkheim E. *Suicide*. New York, NY: Free Press; 1997 [1897].
175. Bowlby J. *Maternal care and mental health*. New York, NY: Schocken; 1951.
176. Sutherland EH. *Principles of criminology*. Chicago, IL: University of Chicago Press; 1924.
177. Rosenstock IM. Why people use health services. *Milbank Mem Fund Q* 1966;**44**:94–127. <http://dx.doi.org/10.1111/j.1468-0009.2005.00425.x>
178. Bandura A. Human agency in social cognitive theory. *Am Psychol* 1989;**44**:1175–84. <http://dx.doi.org/10.1037//0003-066X.44.9.1175>
179. Hirschi T. *Causes of delinquency*. Berkeley, CA: University of California Press; 1969.

180. Bandura A. *Social learning theory*. New York, NY: General Learning Press; 1977.
181. Jencks C, Mayer SE. The social consequences of growing up in a poor neighbourhood. In Lynn L, McGeary M, editors. *Inner-city poverty in the United States*. Washington, DC: National Academy of Sciences; 1990. pp. 111–86.
182. Akers RL, Sellers CS. *Criminological theories: introduction, evaluation, and application*, 4th edn. Los Angeles, CA: Roxbury Publishing Company; 2004.
183. Sherman L. Defiance, deterrence, and irrelevance: a theory of the criminal sanction. *J Res Crime Delinq* 1993;**30**:445–73. <http://dx.doi.org/10.1177/0022427893030004006>
184. Oetting ER, Donnermeyer JF. Primary socialization theory: the etiology of drug use and deviance. I. *Subst Use Misuse* 1991;**33**:995–1026. <http://dx.doi.org/10.3109/10826089809056252>
185. Shaw CR, McKay HD. *Juvenile delinquency in urban areas*. Chicago, IL: University of Chicago Press; 1942.
186. Rotter JB. *Social learning and clinical psychology*. Englewood Cliffs, NJ: Prentice-Hall; 1964.
187. Ajzen I. The theory of planned behavior. In Lange PAM, Kruglanski AW, Higgins ET, editors. *Handbook of theories of social psychology*. Vol. 1. London: Sage; 2012. pp. 438–59.
188. Giddens A. *The constitution of society. Outline of the theory of structuration*. Cambridge: Cambridge University Press; 1984.
189. Mendelson B. A new branch of bio-psychological science: la victimology. *Resume Int Crim Police Tech* 1956;**10**:782–9.
190. Achenbach TM. *Manual for teacher's report form and 1991 profile*. Burlington, MA: University of Vermont, Department of Psychiatry; 1991.

Appendix 1 Searches

TABLE 10 Search (1): MEDLINE core search strategy – 29 July 2010

Set	Searches	Results
1.	health promoting school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	137
2.	(healthy school or healthy schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	448
3.	comprehensive school* health program*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	48
4.	(coordinated school* health program* or co-ordinated school* health program*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	55
5.	1 or 2 or 3 or 4	659
6.	(whole school or school wide or schoolwide).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	207
7.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	165,219
8.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2,479,458
9.	7 and 8	100,116
10.	ethos.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	640
11.	(school* adj5 climate).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	215
12.	exp Climate Change/ or exp Climate/	86,577
13.	11 not 12	195
14.	(school* adj5 environment).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1335
15.	(school* adj5 culture).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	207
16.	(school* adj3 manag*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	556
17.	(school* adj3 leader*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	226
18.	(school* adj5 organi?ation).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	341

continued

TABLE 10 Search (1): MEDLINE core search strategy – 29 July 2010 (continued)

Set	Searches	Results
19.	(school* adj5 aggregat*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	59
20.	(school* adj5 governance).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	27
21.	education* context*.mp.	143
22.	(interschool variation* or inter-school variation*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4
23.	(interschool differen* or inter-school differen*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1
24.	(interschool inequalit* or inter-school inequalit*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	0
25.	(school difference* or differen* between school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	77
26.	(school* adj2 level).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1253
27.	(school* adj3 varia*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	610
28.	(school influence*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	40
29.	((school* adj3 effect) or school) adj3 effects).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	431
30.	restor* justice.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	39
31.	Schools/og [Organization & Administration]	619
32.	Health Promotion/	39,229
33.	31 and 32	89
34.	10 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 33	5808
35.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1,430,864
36.	(wellbeing or well being).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,403
37.	infection*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1,087,334
38.	disease*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3,636,128

TABLE 10 Search (1): MEDLINE core search strategy – 29 July 2010 (*continued*)

Set	Searches	Results
39.	emotion*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	91,977
40.	mental.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	310,841
41.	psychiatr*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	214,115
42.	anxi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	116,132
43.	depress*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	307,948
44.	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43	5,846,882
45.	Juvenile Delinquency/	6383
46.	Violence/pc [Prevention & Control]	3758
47.	(violen* or delinquen* or aggress* or bully* or bullies or bullied).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	145,752
48.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	650,276
49.	victimi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2928
50.	45 or 46 or 47 or 48 or 49	785,497
51.	Alcohol Drinking/	41,336
52.	Smoking/ or Marijuana Smoking/	100,310
53.	Substance-Related Disorders/pc [Prevention & Control]	6277
54.	(smok* or tobacco).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	201,065
55.	cigarett*.mp.	39,495
56.	(illegal drug* or illicit drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	10,677
57.	(drug* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	51,659
58.	(drug* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	22,736
59.	57 not 58	28,923
60.	(drug* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,097
		continued

TABLE 10 Search (1): MEDLINE core search strategy – 29 July 2010 (*continued*)

Set	Searches	Results
61.	(substance* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	12,767
62.	(substance* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1946
63.	61 not 62	10,821
64.	(substance* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	33,382
65.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	14,873
66.	(alcohol or binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	168,866
67.	51 or 52 or 53 or 54 or 55 or 56 or 59 or 60 or 63 or 64 or 65 or 66	416,873
68.	(healthy eating or nutrition or diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	334,194
69.	(obesity or overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	135,458
70.	(bodyweight or body weight or bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	309,228
71.	(physical exercise or physical* activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	45,281
72.	68 or 69 or 70 or 71	681,558
73.	active commuting to school.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	34
74.	(active transport not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2668
75.	(walking bus or walking buses or school travel plan*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1
76.	((walk* adj2 school* or cycl*) adj2 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	145
77.	73 or 74 or 75 or 76	2823
78.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	20,442
79.	exp Contraceptive Devices/ut [Utilization]	3777
80.	Pregnancy in Adolescence/	5869
81.	Sexual Behavior/	34,372
82.	Sexually Transmitted Diseases/ep, pc [Epidemiology, Prevention & Control]	9928

TABLE 10 Search (1): MEDLINE core search strategy – 29 July 2010 (*continued*)

Set	Searches	Results
83.	(pregnan* or sexual).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	795,766
84.	(HIV or chlamydia).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	227,757
85.	(condom* or contracepti*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	71,857
86.	78 or 79 or 80 or 81 or 82 or 83 or 84 or 85	1,038,088
87.	44 or 50 or 67 or 72 or 77 or 86	7,414,552
88.	9 and 34 and 87	2943
89.	6 and 87	173
90.	5 or 88 or 89	3633
91.	schools, dental/ or schools, medical/ or schools, nursing/ or schools, pharmacy/ or schools, public health/ or schools, veterinary/	29,874
92.	medical school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	19,144
93.	91 or 92	43,387
94.	90 not 93	3518

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010

Set	Searches	Results
1.	health promoting school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	139
2.	(healthy school or healthy schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	450
3.	comprehensive school* health program*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	48
4.	(coordinated school* health program* or co-ordinated school* health program*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	56
5.	1 or 2 or 3 or 4	664
6.	(whole school or school wide or schoolwide).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	208
7.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	165,694
		<i>continued</i>

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
8.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2,485,450
9.	7 and 8	100,468
10.	ethos.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	641
11.	(school* adj5 climate).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	218
12.	exp Climate Change/ or exp Climate/	86,839
13.	11 not 12	198
14.	(school* adj5 environment).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1339
15.	(school* adj5 culture).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	207
16.	(school* adj3 manag*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	558
17.	(school* adj3 leader*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	228
18.	(school* adj5 organi?ation).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	341
19.	(school* adj5 aggregat*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	60
20.	(school* adj5 governance).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	27
21.	education* context*.mp.	145
22.	(interschool variation* or inter-school variation*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4
23.	(interschool differen* or inter-school differen*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1
24.	(interschool inequalit* or inter-school inequalit*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	0
25.	(school difference* or differen* between school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	78
26.	(school* adj2 level).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1262

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
27.	(school* adj3 varia*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	611
28.	(school influence*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	40
29.	((school* adj3 effect) or school) adj3 effects).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	438
30.	restor* justice.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	40
31.	Schools/og [Organization & Administration]	621
32.	Health Promotion/	39,392
33.	31 and 32	89
34.	10 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 33	5841
35.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1,435,980
36.	(wellbeing or well being).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,542
37.	infection*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1,090,528
38.	disease*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3,646,240
39.	emotion*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	92,408
40.	mental.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	311,800
41.	psychiatr*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	214,674
42.	anxi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	116,609
43.	depress*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	308,909
44.	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43	5,863,435
45.	Juvenile Delinquency/	6409
46.	Violence/pc [Prevention & Control]	3768
		continued

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
47.	(violen* or delinquen* or aggress* or bully* or bullies or bullied).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	146,311
48.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	652,077
49.	victimi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2958
50.	45 or 46 or 47 or 48 or 49	787,823
51.	Alcohol Drinking/	41,502
52.	Smoking/ or Marijuana Smoking/	100,634
53.	Substance-Related Disorders/pc [Prevention & Control]	6299
54.	(smok* or tobacco).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	201,762
55.	cigarett*.mp	39,637
56.	(illegal drug* or illicit drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	10,721
57.	(drug* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	51,908
58.	(drug* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	22,831
59.	57 not 58	29,077
60.	(drug* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,189
61.	(substance* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	12,891
62.	(substance* (adj2 used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1949
63.	61 not 62	10,942
64.	(substance* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	33,519
65.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	14,930
66.	(alcohol or binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	169,462
67.	51 or 52 or 53 or 54 or 55 or 56 or 59 or 60 or 63 or 64 or 65 or 66	418,331
68.	(healthy eating or nutrition or diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	335,062

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
69.	(obesity or overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	136,129
70.	(bodyweight or body weight or bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	310,258
71.	(physical exercise or physical* activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	45,514
72.	68 or 69 or 70 or 71	683,716
73.	active commuting to school.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	34
74.	(active transport not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2671
75.	(walking bus or walking buses or school travel plan*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1
76.	((walk* adj2 school*) (or cycl*) adj2 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	145
77.	73 or 74 or 75 or 76	2826
78.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	20,453
79.	exp Contraceptive Devices/ut [Utilization]	3800
80.	Pregnancy in Adolescence/	5879
81.	Sexual Behavior/	34,470
82.	Sexually Transmitted Diseases/ep, pc [Epidemiology, Prevention & Control]	9976
83.	(pregnan* or sexual).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	797,455
84.	(HIV or chlamydia).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	228,565
85.	(condom* or contracepti*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	72,004
86.	78 or 79 or 80 or 81 or 82 or 83 or 84 or 85	1,040,475
87.	44 or 50 or 67 or 72 or 77 or 86	7,434,634
88.	9 and 34 and 87	2966
89.	6 and 87	174
		<i>continued</i>

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
90.	5 or 88 or 89	3660
91.	schools, dental/ or schools, medical/ or schools, nursing/ or schools, pharmacy/ or schools, public health/ or schools, veterinary/	29,917
92.	medical school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	19,187
93.	91 or 92	43,464
94.	90 not 93	3545
95.	(multiintervention or multi-intervention).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	30
96.	(noncurricul* or non-curric*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	15
97.	(socioecolog* or socio-ecolog*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	264
98.	(socioenvironment* or socio-environment*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	854
99.	classroom management.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	44
100.	(value added not (tax or VAT)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	794
101.	pastoral.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3566
102.	((school* adj3 achievement*) or (school* adj3 attainment*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1059
103.	((school* adj3 exam*) or (school adj3 (test or tests or testing or tested))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2883
104.	(ecological level or ecological study or ecological studies or ecological design* or ecological measure* or ecological variable* or ecological analysis or school ecology).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2368
105.	((student* engagement or engaging students or pupil* engagement or engaging pupils or engaging school or (school* adj2 engagement) or (education* adj2 engagement)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	235
106.	95 or 96 or 97 or 98 or 99 or 100 or 101 or 104 or 105	8147
107.	((school* adj3 qualif*) or (school* adj3 quality) or (school* adj3 inspect*) or (school* adj3 influence*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1125

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
108.	((school* adj5 (policy or policies)) or (school* adj3 rules) or (school* adj3 practices)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1648
109.	((school* adj5 context*) or (school* adj3 opportunit*) or (school* adj5 collective)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	787
110.	((school* adj3 communit*) or (school* adj3 structur*) or (school* adj3 relation*) or (school* adj3 communicat*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3437
111.	((school* adj3 structur*) or (school* adj3 security) or (school* adj3 safe*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	590
112.	((school* adj3 expectation*) or (school* adj2 exclusion*) or (school inclusion or between adj2 schools)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	661
113.	102 or 103 or 107 or 108 or 109 or 110 or 111 or 112	11,145
114.	((education* adj3 achievement*) or (education* adj3 attainment*) or (education* adj3 exam*) or (education* adj3 (tests or test or testing))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	6087
115.	((education* adj3 qualif*) or (education* adj3 quality) or (education* adj3 (policy or policies.mp))). [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4058
116.	((education* adj3 opportunit*) or education* adj3practices or (education* adj3 culture) or (education* adj3 manag*) or (education* adj3 leader*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	5904
117.	((education* adj3 practices) or (education* adj3 communicat*) or (education* adj3 safe*) or (education* adj3 expectation*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3730
118.	114 or 115 or 116 or 117	19,103
119.	((teaching adj3 practices) or (teaching adj3 standard*) or (teaching adj3 style*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	895
120.	((teaching adj3 method*) or (teaching adj3 differen*) or (teaching adj3 varia*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4857
121.	119 or 120	5643

continued

TABLE 11 Search (2): MEDLINE non-core search strategy – 12 August 2010 (*continued*)

Set	Searches	Results
122.	(pupil-led or student led).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	53
123.	aggregate* adj2 data or reports or information.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1591
124.	(school* size or school restructur* or comprehensive school reform).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	72
125.	122 or 123 or 124	1716
126.	106 or 113 or 118 or 121 or 125	44,608
127.	9 and 87 and 126	8235
128.	127 not 93	7767
129.	128 not 94	6793

TABLE 12 Search (3): MEDLINE additional search strategy – 23 September 2010

Set	Searches	Results
1.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	2,568,402
2.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	167,005
3.	1 and 2	113,129
4.	((active transport adj3 school*) not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	13
5.	active commuting to school.mp.. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	34
6.	(walking bus or walking buses or walking school bus*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	7
7.	school travel plan*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1
8.	(walk* adj3 school*).ti.ab.	130
9.	((cycle adj3 school*) or (cycling adj3 school*) or (bicycl* adj3 school*)).ti.ab.	85
10.	4 or 5 or 6 or 7 or 8 or 9	228
11.	3 and 10	213

TABLE 12 Search (3): MEDLINE additional search strategy – 23 September 2010 (*continued*)

Set	Searches	Results
12.	((school* adj3 dinner*) or (school* adj3 lunch*) or (school* adj3 breakfast*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	646
13.	breakfast club*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	8
14.	(soft drink* machine* or snack* machine*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	6
15.	12 or 13 or 14	655
16.	3 and 15	595
17.	school meals.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	234
18.	(school* adj3 meal*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	354
19.	18 not 17	120
20.	3 and 19	111
21.	16 or 20	671
22.	physical train*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	3762
23.	"Physical Education and Training"/	10,602
24.	(sport* adj3 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	512
25.	(games adj3 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	28
26.	22 or 23 or 24 or 25	13,770
27.	3 and 26	1462
28.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1,448,176
29.	(wellbeing or well-being).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,911
30.	(depress* or emotion* or mental or psychiatr* or anxi*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	811,688
31.	(infection* or disease*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4,298,493
32.	28 or 29 or 30 or 31	5,905,334
33.	Violence/pc [Prevention & Control]	3799
		continued

TABLE 12 Search (3): MEDLINE additional search strategy – 23 September 2010 (*continued*)

Set	Searches	Results
34.	(bully* or bullied or bullies).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1193
35.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	656,888
36.	Alcohol Drinking/	41,859
37.	Substance-Related Disorders/pc [Prevention & Control]	6337
38.	Smoking/ or Marijuana Smoking/	101,457
39.	(substance* adj2 (user* or abuse* or misuse*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	33,547
40.	(smok* or tobacco).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	203,627
41.	cigarett*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	39,976
42.	(drug* adj2 (user* or abuse* or misuse*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	25,626
43.	(illicit drug* or illegal drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	7310
44.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	15,055
45.	(alcohol).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	167,916
46.	(binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	4669
47.	(healthy eating or healthy-eating).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	1388
48.	(obese or obesity).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	142,720
49.	(nutrition).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	118,291
50.	(diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	241,314
51.	(overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	23,296
52.	(bodyweight or body weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	229,595
53.	(bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	101,322

TABLE 12 Search (3): MEDLINE additional search strategy – 23 September 2010 (*continued*)

Set	Searches	Results
54.	(physical exercise or physical* activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	46,166
55.	47 or 48 or 49 or 50 or 51 or 52 or 53 or 54	696,215
56.	Condoms/ut [Utilization]	3065
57.	HIV Infections/ep, pc [Epidemiology, Prevention & Control]	36,132
58.	teenage pregnancy.mp. or Pregnancy in Adolescence/	6203
59.	Sexual Behavior/	34,770
60.	Sexually Transmitted Diseases/ep, pc [Epidemiology, Prevention & Control]	10,056
61.	sexual.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	137,750
62.	pregnan*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	678,875
63.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	20,524
64.	HIV*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	211,093
65.	chlamydia*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	20,869
66.	(condom or condoms or contracept*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	72,313
67.	56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66	1,046,418
68.	(violen* or aggress* or victimi*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	142,137
69.	33 or 34 or 35 or 68	787,707
70.	((drug* adj2 use?) not (drug* adj2 (used or uses))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	29,411
71.	((substance* adj2 use?)not (substance* adj2 (used or uses))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	11,154
72.	36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 70 or 71	419,192
73.	32 or 55 or 67 or 69 or 72	7,482,875
74.	27 and 73	1144
75.	11 and 73	183
76.	21 and 73	576

Appendix 2 Included studies/theory for stage 2 in-depth review

Research question 1: theories

Bonding theory/social bonding theory

Contagion theory/model

Control theory/social control theory

Defiance theory

Deterrence theory

Differential association theory

Ecological model of co-ordinated school health programmes

Ecological systems theory

Human functioning and school organisation

Integrated perspective on delinquent behaviour

Moral authority theory

Primary socialisation theory

Problem behavior theory

Social capital theory

Social cognitive theory

Social development model/theory

Social disorganisation/integration theory

Social learning theory

Strain theory

Theory of planned behaviour

Theory of reasoned action

Theory of structuration

Theory of triadic influence

Victimisation theory

Research question 2: outcome evaluation studies

Battistich V. Effects of a school-based program to enhance prosocial development on children's peer relations and social adjustment. *J Res Character Educ* 2003;**1**:1–7.

Battistich V, Solomon D, Delucchi KL. Effects of a program to enhance prosocial development on adjustment. Annual Meeting of the American Psychological Association, Boston, MA, 10–14 August 1990.

Battistich V, Solomon D, Watson M, Solomon J, Schaps E. Effects of an elementary school program to enhance prosocial behavior on children's cognitive-social problem solving skills and strategies. *J Appl Dev Psychol* 1989;**10**:147–69.

Bonell C, Sorhaindo A, Allen ES, Strange VJ, Wiggins M, Fletcher A, *et al*. Pilot multimethod trial of a school-ethos intervention to reduce substance use: building hypotheses about upstream pathways to prevention. *J Adolesc Health* 2010;**47**:555–63.

Dzewaltowski DA, Estabrooks PA, Welk G, Hill J, Milliken G, Karteroliotis K, *et al*. Healthy youth places: a randomized controlled trial to determine the effectiveness of facilitating adult and youth leaders to promote physical activity and fruit and vegetable consumption in middle schools. *Health Educ Behav* 2009;**36**:583–600.

Flannery DJ, Vazsonyi AT, Liau AK, Guo S, Powell KE, Atha H, *et al*. Initial behavior outcomes for the peacebuilders universal school-based violence prevention program. *Dev Psychol* 2003;**39**:292–308.

Flay BR, Graumlich S, Segawa E, Burns JL, Holliday MY. Effects of 2 prevention programs on high-risk behaviors among African American youth: a randomized trial. *Arch Pediatr Adolesc Med* 2004;**158**:377–84.

Fonagy P, Twemlow SW, Vernberg EM, Nelson JM, Dill EJ, Little TD, *et al*. A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *J Child Psychol Psychiatry* 2009;**50**:607–16.

Krug EG, Brener ND, Dahlberg LL, Ryan GW, Powell KE. The impact of an elementary school-based violence prevention program on visits to the school nurse. *Am J Prev Med* 1997;**13**:459–63.

Ridgers ND, Fairclough SJ, Stratton G. Twelve-month effects of a playground intervention on children's morning and lunchtime recess physical activity levels. *J Phys Act Health* 2010;**7**:167–75.

Ridgers ND, Stratton G, Fairclough SJ, Twisk JWR. Long-term effects of a playground marking and physical structures on children's recess physical activity levels. *Prev Med* 2007;**44**:393–7.

Ridgers ND, Stratton GS, Fairclough SJ, Twisk JWR. Children's physical activity levels during school recess: a quasi-experimental intervention study. *Int J Behav Nutr Phys Act* 2007;**4**:1–9.

Sallis JF, McKenzie TL, Conway TL, Elder JP, Prochaska JJ, Brown M, *et al*. Environmental interventions for eating and physical activity: a randomized controlled trial in middle schools. *Am J Prev Med* 2003;**24**:209–17.

Solomon D, Watson MS, Delucchi KL, Schaps E, Battistich V. Enhancing children's prosocial behavior in the classroom. *Am Educ Res J* 1988;**25**:527–54.

Stratton G. Promoting children's physical activity in primary school: an intervention study using playground markings. *Ergonomics* 2000;**43**:1538–46.

Stratton G, Mullan E. The effect of multicolor playground markings on children's physical activity level during recess. *Prev Med* 2005;**41**:828–33.

Research question 3: process evaluation studies

Battistich V, Solomon D, Watson M, Solomon J, Schaps E. Effects of an elementary school program to enhance prosocial behavior on children's cognitive-social problem solving skills and strategies. *J Appl Dev Psychol* 1989;**10**:147–69.

Bonell C, Sorhaindo A, Allen ES, Strange VJ, Wiggins M, Fletcher A, et al. Pilot multimethod trial of a school-ethos intervention to reduce substance use: building hypotheses about upstream pathways to prevention. *J Adolesc Health* 2010;**47**:555–63.

Bonell C, Sorhaindo A, Strange VJ, Wiggins M, Allen E, Fletcher A, et al. A pilot whole-school intervention to improve school ethos and reduce substance use. *Health Educ* 2010;**110**:252–72.

Dzewaltowski DA, Estabrooks PA, Welk G, Hill J, Milliken G, Karteroliotis K, et al. Healthy youth places: a randomized controlled trial to determine the effectiveness of facilitating adult and youth leaders to promote physical activity and fruit and vegetable consumption in middle schools. *Health Educ Behav* 2009;**36**:583–600.

Flannery DJ, Vazsonyi AT, Liao AK, Guo S, Powell KE, Atha H, et al. Initial behavior outcomes for the peacebuilders universal school-based violence prevention program. *Dev Psychol* 2003;**39**:292–308.

Solomon D, Watson MS, Delucchi KL, Schaps E, Battistich V. Enhancing children's prosocial behavior in the classroom. *Am Educ Res J* 1988;**25**:527–54.

Research question 4: multilevel studies

Adams ML, Jason LA, Pokorny S, Hunt Y. The relationship between school policies and youth tobacco use. *J School Health* 2009;**79**:17–23.

Akiba M. What predicts fear of school violence among us adolescents? *Teach Coll Rec* 2010;**112**:68–102.

Attar-Schwartz S, Khoury-Kassabri M. Indirect versus verbal forms of victimization at school: the contribution of student, family, and school variables. *Soc Work Res* 2008;**32**:159–70.

Aveyard A, Markham WA, Chenge KK. A methodological and substantive review of the evidence that schools cause pupils to smoke. *Soc Sci Med* 2004;**58**:2252–65.

Bisset S, Markham WA, Aveyard P. School culture as an influencing factor on youth substance use. *J Epidemiol Community Health* 2007;**61**:485–90.

Bradshaw CP, Sawyer AL, O'Brennan LM. A social disorganization perspective on bullying-related attitudes and behaviors: the influence of school context. *Am J Community Psychol* 2009;**43**:204–20.

Collins D, Pan Z, Johnson K, Courser M, Shamblen S. Individual and contextual predictors of inhalant use among 8th graders: a multilevel analysis. *J Drug Educ* 2008;**38**:193–210.

Comeau C. *Impact of school context on the relations between deviant peer affiliations and problem behaviors during the middle school years: an exploratory analysis using hierarchical linear modeling*. PhD thesis. Eugene, OR: University of Oregon; 2005.

Cradock AL, Melly SJ, Allen JG, Morris JS, Gortmaker SL. Characteristics of school campuses and physical activity among youth. *Am J Prev Med* 2007;**33**:106–13.

Evans-Whipp T, Beyers JM, Lloyd S, Lafazia AN, Toumbourou JW, Arthur MW, et al. A review of school drug policies and their impact on youth substance use. *Health Promot Int* 2004;**19**:227–34.

- Gastic B. *Student attitudes about school safety and their implications for safe school reform*. PhD thesis. Stanford, CA: Stanford University; 2005.
- Gladden RM. *Building school communities and making schools safer*. PhD thesis. New York, NY: City University of New York; 2005.
- Grunseit AC, Weatherburn D, Donnelly N. Correlates of physical violence at school: a multilevel analysis of Australian high school students. *Aust J Soc Issues* 2008;**43**:527–45.
- Haley SJ. *School context and the association of risk and protective factors among same-sex attracted youth*. Dissertation. Waltham, MA: Brandeis University; 2007.
- Haug E, Torsheim T, Samdal O. Local school policies increase physical activity in Norwegian secondary schools. *Health Promot Int* 2010;**25**:63–72.
- Henry KL, Stanley LR, Edwards RW, Harkabus LC, Chapin LA. Individual and contextual effects of school adjustment on adolescent alcohol use. *Prev Sci* 2009;**10**:236–47.
- Ishibashi SLW. *Lessons learned from tobacco control: a multilevel analysis of school characteristics and adolescent physical activity*. PhD thesis. Waterloo, ON: University of Waterloo; 2009.
- Johnson R, Hoffmann J. Adolescent cigarette smoking in US racial/ethnic subgroups: findings from a national longitudinal study. *J Health Soc Behav* 2000;**41**:392–407.
- Kairouz S, Adlaf EM. Schools, students and heavy drinking: a multilevel analysis. *Addict Res Theory* 2003;**11**:427–39.
- Khoury-Kassabri M, Astor RA, Benbenishty R. Middle Eastern adolescents' perpetration of school violence against peers and teachers: a cross-cultural and ecological analysis. *J Interpers Violence* 2009;**24**:159–82.
- Kim S. *Prevention of adolescent substance use: an investigation of cross-level interaction effects of self, peer, family and school level risk and protective factors using multilevel modeling*. PhD thesis. Los Angeles, CA: University of Southern California; 1999.
- Kim YS. *The effects of individual, neighborhood, and school characteristics on delinquency: a multi-level cross-nested analysis*. PhD thesis. University Park, PA: Pennsylvania State University; 2001.
- Kumar R, O'Malley PM, Johnston LD. Association between physical environment of secondary schools and student problem behavior: a national study, 2000–2003. *Environ Behav* 2008;**40**:455–86.
- Leatherdale ST, Manske S, Faulkner G, Arbour K, Bredin C. A multi-level examination of school programs, policies and resources associated with physical activity among elementary school youth in the play-on study. *Int J Behav Nutr Phys Act* 2010;**7**:1–13.
- Lothian SF. *Understanding student smoking behavior: an ecological approach*. PhD thesis. London, ON: University of Western Ontario; 2003.
- Ma X. Health outcomes of elementary school students in New Brunswick. The education perspective. *Eval Rev* 2000;**24**:435–56.
- Maes L, Lievens J. Adolescent smoking behaviour and the school: a multi-level analysis. In Tudor-Smith C, editor. *Tackling tobacco*. Cardiff: Health Promotion Wales; 1999. pp. 113–28.
- Malikaew P, Watt RG, Sheiham A. Associations between school environments and childhood traumatic dental injuries. *Oral Health Prev Dent* 2003;**1**:255–66.
- Markham WA, Aveyard P, Bisset SL, Lancashire ER, Bridle C, Deakin S. Value-added education and smoking uptake in schools: a cohort study. *Addiction* 2008;**103**:155–61.
- Monshouwer K, Van Dorsselaer S, Van Os J, Drukker M, De Graaf R, Ter Bogt T, et al. Ethnic composition of schools affects episodic heavy drinking only in ethnic-minority students. *Addiction* 2007;**102**:722–9.

- Moore L, Roberts C, Tudor-Smith C. School smoking policies and smoking prevalence among adolescents: a multi-level analysis of cross-sectional data from Wales. *Tob Control* 2001;**10**:117–23.
- Murnaghan DA, Sihvonen M, Leatherdale ST, Kekki P. The relationship between school-based smoking policies and prevention programs on smoking behavior among grade 12 students in Prince Edward Island: a multilevel analysis. *Prev Med* 2007;**44**:317–22.
- Pinilla J, González B, Barber P, Santana Y. Smoking in young adolescents: an approach with multilevel discrete choice models. *J Epidemiol Community Health* 2002;**56**:227–32.
- Piontek D, Buehler A, Donath C, Floeter S, Rudolph U, Metz K, et al. School context variables and students' smoking. *Eur Addict Res* 2008;**14**:53–60.
- Rountree PW, Clayton RR. A contextual model of adolescent alcohol use across the rural–urban continuum. *Subst Use Misuse* 1999;**34**:495–519.
- Tobler AL, Komro KA, Dabroski A, Aveyard P, Markham WA. Preventing the link between SES and high-risk behaviors: 'value-added' education, drug use and delinquency in high-risk, urban schools. *Prev Sci* 2011;**12**:211–21.
- van der Horst K, Oenema A, Velde T, Brug J. Gender, ethnic and school type differences in overweight and energy balance-related behaviours among Dutch adolescents. *Int J Pediatr Obes* 2009;**4**:371–80.
- Way SM. *For their own good? The effects of school discipline and disorder on student behavior and academic achievement*. Dissertation. Tucson, AZ: University of Arizona; 2003.
- West P, Sweeting H, Leyland A. School effects on pupils health behaviours: evidence in support of the health promoting school. *Res Papers Educ* 2004;**19**:261–91.
- West P, Sweeting H, Young R. Transition matters: pupils' experiences of the primary–secondary school transition in the west of Scotland and consequences for well-being and attainment. *Res Papers Educ* 2010;**25**:21–50.
- Wiium N, Wold B. Family and school influences on adolescent smoking behaviour. *Health Educ* 2006;**106**:465–79.
- Wiium N, Wold B. An ecological system approach to adolescent smoking behavior. *J Youth Adolesc* 2009;**38**:1351–63.

Research question 5: qualitative studies

- Astor RA, Meyer HA, Behre WJ. Unowned places and times: maps and interviews about violence in high schools. *Am Educ Res J* 1999;**36**:3–42.
- Astor RA, Meyer HA, Pitner RO. Elementary and middle school students' perceptions of violence-prone school subcontexts. *Elem Sch J* 2001;**101**:511–28.
- Brotman JS. *Urban high school students' talk about HIV/AIDS decision-making: learning, identities, and the influence of school*. PhD thesis. New York, NY: Columbia University; 2009.
- Brunson RK, Miller JM. Schools, neighborhoods, and adolescent conflicts: a situational examination of reciprocal dynamics. *Justice Q* 2009;**26**:183–210.
- Burnett C. School violence in an impoverished South African community. *Child Abuse Negl* 1998;**22**:789–95.
- Cousins LH. Toward a sociocultural context for understanding violence and disruption in black urban schools and communities. *J Sociol Soc Welf* 1997;**24**:41–63.
- Devine J. Can metal detectors replace the panopticon? *Cult Anthropol* 1995;**10**:171–95.

- Fletcher A, Bonell C, Rhodes T. New counter-school cultures: female students' drug use at a high-achieving secondary school. *Br J Sociol Educ* 2009;**30**:549–62.
- Fletcher A, Bonell C, Sorhaindo A, Rhodes T. Cannabis use and 'safe' identities in an inner-city school risk environment. *Int J Drug Policy* 2009;**20**:244–50.
- Fletcher A, Bonell C, Sorhaindo A, Strange V. How might schools influence young people's drug use? Development of theory from qualitative case-study research. *J Adolesc Health* 2009;**45**:126–32.
- Gordon J, Turner K. School staff as exemplars – where is the potential? *Health Educ* 2001;**101**:283–91.
- Haselswerdt MV, Lenhardt AMC. Reframing school violence: listening to voices of students. *Educ Forum* 2003;**67**:326–36.
- Hosie ACS. 'I hated everything about school': an examination of the relationship between dislike of school, teenage pregnancy and educational disengagement. *Soc Policy Soc* 2007;**6**:333–47.
- Keddie A. 'Some of those girls can be real drama queens': issues of gender, sexual harassment and schooling. *Sex Educ* 2009;**9**:1–16.
- Lundblad B, Hellstrom AL, Berg M. Children's experiences of attitudes and rules for going to the toilet in school. *Scand J Caring Sci* 2010;**24**:219–23.
- Marsiglia FF, Miles BW, Dustman P, Sills S. Ties that protect: an ecological perspective on Latino/a urban pre-adolescent drug use. *J Ethn Cult Divers Soc Work* 2002;**11**:191–220.
- Pike J, Colquhoun D. The relationship between policy and place: the role of school meals in addressing health inequalities. *Health Sociol Rev* 2009;**18**:50–60.
- Plano Clark VL, Miller DL, Creswell JW, McVea K, McEntarffer R, Harter LM, *et al*. In conversation: high school students talk to students about tobacco use and prevention strategies. *Qual Health Res* 2002;**12**:1264–83.
- Waldron LM. The messy nature of discipline and zero tolerance policies: negotiating safe school environments among inconsistencies, structural constraints and the complex lives of youth. *Sociol Stud Child Youth* 2005;**11**:81–114.
- Wills W, Backett-Milburn K, Gregory S, Lawton J. The influence of the secondary school setting on the food practices of young teenagers from disadvantaged backgrounds in Scotland. *Health Educ Res* 2005;**20**:458–65.
- Wilson-Simmons R, Dash K, Tehranifar P, O'Donnell L, Stueve A. What can student bystanders do to prevent school violence? Perceptions of students and school staff. *J Sch Violence* 2006;**5**:43–62.

Appendix 3 Stage 1 evidence and theory map countries listed in included references

TABLE 13 Countries of research of references included in the evidence and theory map stage 1

Country of research	<i>n</i>
Antigua and Barbuda	1
Australia	74
Austria	1
Belgium	5
Brazil	8
Bulgaria	2
Canada	24
Chile	3
China	14
Columbia	3
Costa Rica	1
Croatia	1
Denmark	6
Ecuador	1
Egypt	1
Finland	13
France	2
Germany	11
Greece	4
Holland	1
Hong Kong	10
India	1
Indonesia	2
Ireland	1
Israel	7
Italy	4
Kenya	3
Republic of Korea	2
Kuwait	1
Lao People's Democratic Republic	1
Latvia	1
Malaysia	2

continued

TABLE 13 Countries of research of references included in the evidence and theory map stage 1 (continued)

Country of research	<i>n</i>
Mauritius	1
Mexico	6
Myanmar	1
Netherlands	17
New Zealand	15
Nigeria	2
Norway	23
Pakistan	1
Panama	1
Papua New Guinea	1
Peru	3
Philippines	1
Poland	4
Portugal	1
Russian Federation	3
Serbia	1
Slovakia	1
South Africa	11
Republic of Korea	2
Spain	7
Sweden	5
Switzerland	7
Taiwan, Province of China	10
United Republic of Tanzania	1
Thailand	8
Trinidad and Tobago	1
Turkey	3
Uganda	3
UK	99
USA	322
Various Latin American countries	2
Yugoslavia	1
Zimbabwe	1

Appendix 4 Stage 1 theory map

TABLE 14 Theory map stage 1

Name of theory	Summary	Possible pathways	Health outcomes that this suggests might be important	Queries and critique
Anomie theory Durkheim 1997 [1897] ¹⁷⁴	Anomie means 'personal feeling of a lack of norms'. It arises from a mismatch between personal values and wider social values, or from the lack of a social ethic, producing moral deregulation and absence of legitimate aspirations. Anomie occurs when society has undergone significant historical changes and when there is a discrepancy between societal values and what individuals can actually achieve	The school succeeds in setting standards and communicates norms and values, and these mesh with those already held by students and/or by society in general. The school instils aspirations that fit with student values and are achievable. Staff behaviour matches the school's espoused values. The school helps students develop communal values and aspirations. Students avoid behaviour that is antisocial or which does not contribute to the achievement of future goals. Students have realistic goals that they can achieve legitimately and therefore avoid ASB and defer immediate gratification	Outcomes concerning ASB (under-16 sex or smoking, under-18 drinking, illicit drug use and violence), immediate gratification (teenage pregnancy, diet, exercise) and egoistic values (emotional/mental health, eating/body disorders and violence)	Its application to school health assumes that schools can make a difference without addressing broader societal anomie
Attachment theory Bowlby 1951 ¹⁷⁵	A young child needs to develop a relationship with at least one primary caregiver for normal social and emotional development. Infants seek proximity to an attachment figure in stressful situations. Infants become attached to adults who are sensitive and responsive, and who remain as consistent caregivers. Attachments to peers tend to emerge in adolescence, although parents continue to be attachment figures. With adolescents, the role of the parental figures is to be available when needed while the adolescent makes excursions into the outside world	School ensures that students have at least one long-term relationship with a sensitive and responsive staff member providing pastoral care in regular and close contact, and that the staff member is available when the student encounters stressful situations. School ensures that it co-ordinates with parents in providing support to young people. School attachment encourages personal development and emotional health. In turn, this fosters self-preserving rather than risky behaviours	All outcomes, but especially emotional/mental health outcomes	But is attachment theory generalisable from infants to adolescents? Is positive development inconsistent with some engagement with risk?

Name of theory	Summary	Possible pathways	Health outcomes that this suggests might be important	Queries and critique
Differential association theory	Criminal behaviour is learned in interaction with others, especially within intimate personal groups. Learning includes techniques of committing the crime, and the motives, drives, rationalisations and attitudes associated with crime. A person becomes delinquent because of an excess of definitions favourable to violation of law over definitions unfavourable to this. Although criminal behaviour is an expression of general needs and values it is not explained by those needs and values, as non-criminal behaviour is an expression of the same needs and values	School reduces contacts between students engaging/not engaging in criminal or ASBs through fixed and permanent exclusions, internal exclusions and streaming. School ensures that students are exposed to antiracial messages. Students are discouraged from wanting to or being able to engage in crime and ASBs	ASB (under-16 sex or smoking, under-18 drinking, illicit drug use, violence)	Crime is primarily a learned activity. But won't some criminal and ASBs not be learnt but arise from broader developmental or environmental problems?
Ecological systems theory Bronfenbrenner 1979 ¹²	Human development involves accommodation between a growing individual and his or her immediate settings, affected by relations between these settings and the larger contexts in which the settings are embedded. The microsystem concerns the individual's activities, roles and relations in a given setting. The mesosystem is the inter-relations among these settings e.g. the relations among home, school and neighbourhood peer group. The exosystem is the settings that do not involve the developing person but which affect him or her. The macrosystem is the consistencies, in the form and context of lower-order systems plus the belief systems underlying these. Each system contains roles, norms and rules that may shape psychological development	The school allows students to take on roles that are perceived as positive by them. The school communicates norms and rules regarding students' conduct, and norms regarding students' educational engagement and expectations. Students' active participation in lessons and the wider school community as well as their acceptance of clear rules and norms encourages them to develop a positive sense of self, positive educational engagement and expectations and prosocial attitudes and behaviours	All but especially ASB (under-16 sex or smoking, under-18 drinking, illicit drug use, violence) and emotional/mental health outcomes	Can schools offer young people roles that fit with their roles in other systems such as the family and peers? But what if these roles contradict each other (e.g. students have adult-like responsibilities elsewhere)? Do school rules and norms always support positive development? But don't some aspects of school hamper development for some students (e.g. labelling them as unintelligent or restricting their engagement in non-academic activities)?

continued

TABLE 14 Theory map stage 1 (continued)

Name of theory	Summary	Possible pathways	Health outcomes that this suggests might be important	Queries and critique
Educational transmission of class theory Bernstein 1975 ¹⁰⁵	Schools relay two inter-related yet separate orders: the instructional order lessons and the regulatory order discipline. Students respond, becoming committed, alienated, detached or estranged	Schools engage young people in learning and motivate positive development. Schools successfully instil discipline, which encourages prosocial development. Students feel connected with school, motivated to avoid behaviours that are regarded as anti-school and drawn towards behaviours that facilitate personal development	All but especially antisocial outcomes (under-16 sex or smoking, under-18 drinking, illicit drug use, violence) and outcomes concerning aspirations (teenage pregnancy) and emotional/mental health	But do schools operate instructional and regulatory orders that are the same (fair) to all students or do they discriminate?
Health belief model Rosenstock 1966 ¹⁷⁷	The likelihood of an individual undertaking a health action is determined by perceived susceptibility, severity, barriers and costs. Later were added perceived efficacy, cues to action and health motivation	School sets rules about, and punishes, behaviours such as smoking, drinking, taking drugs and violence. School prevents students from accessing unhealthy food. School rewards sporting and other physical activity, and encourages active transport. School educates students about their susceptibility to health risks, and the severity of the costs associated with these, and provides skills in implementing healthy behaviours. These punishments and rewards encourage healthy behaviours without necessarily changing underlying beliefs about the behaviours	All	Individuals rationally weigh up multiple factors before making a decision. But aren't they also unconsciously influenced by peers or habit?
Social cognitive theory Bandura 1989 ¹⁷⁸	Builds on social learning theory. People learn by observing others, with the environment, behaviour and cognition all being important in development. These three factors are reciprocally inter-related (e.g. behaviour witnessed can change cognition, parent's cognition determines the environment that children are raised in)	As per social learning theory (see below) – but the pathways will be influenced and mediated by student cognitive factors	As per social learning theory	Students will learn from social superiors (teachers) and role models (peers) but in cognitively interpreting this will assess what they think about that person and the behaviour – so teachers may not be influential

Name of theory	Summary	Possible pathways	Health outcomes that this suggests might be important	Queries and critique
Social control theory Hirschi 1969 ¹⁷⁹	Social control operates through rules about how people should or should not behave. Informal social control can be given by families, schools, etc., and formal social control can be given by the courts and police. Those who commit delinquent acts do so because ties to the conventional order are weakened. Social control theory derives from the functionalist theory of crime	Schools focus on students' social development and on instilling norms of prosocial behaviour. Teachers model prosocial behaviours. Connection with school leads to prosocial not antisocial attitudes, and self-control, which leads to reductions in ASB	Antisocial outcomes (under-16 sex or smoking, under-18 drinking, illicit drug use, violence)	Are conventional norms really always prosocial (e.g. focus on individual achievement, materialism and competitiveness)? Do institutions sometimes instil questionable values?
Social development model Hawkins and Weiss 1985 ⁹⁸	Providing individuals with opportunities, skills and recognition encourages bonding (attachment and commitment), which encourages healthy beliefs and clear standards, which in turn promotes healthy behaviours	Schools engage all students in learning by providing interesting lessons, teaching useful skills and recognising achievement. Engagement with learning encourages students to develop high expectations and belief in prosocial standards of behaviour, which leads to reductions in delinquency and behaviours that threaten future expectations	All but especially antisocial outcomes (under-16 sex or smoking, under-18 drinking, illicit drug use, violence) and emotional/mental health outcomes	What is the mechanism by which bonding to school encourages healthy beliefs and behaviours – acceptance of norms? Focus on future success rather than immediate pleasure?
Social learning theory Bandura 1977 ¹⁸⁰	People learn through observing/imitating others, assessing whether behaviour is associated with desired outcomes or with social superiors/role models. Social learning requires retention (remembering what one observed), reproduction (ability to reproduce the behaviour) and motivation (good reason) to want to adopt the behaviour	Staff have good relationships with students and model healthy behaviours. Staff communicate the rationale for healthy behaviours. Students are given opportunities to imitate these behaviours in school and are rewarded for this. Students then model positive behaviours to each other	All but especially public behaviours (e.g. healthy eating, physical activity, violence)	Students might regard staff as social superiors or role models. But won't they actually be seen as socially distant, especially by students from different backgrounds?

continued

TABLE 14 Theory map stage 1 (continued)

Name of theory	Summary	Possible pathways	Health outcomes that this suggests might be important	Queries and critique
Strain theory Merton 1938 ¹¹⁰	Informed by theories of anomie. Social structures may encourage individuals to commit crimes through the discrepancy between common social goals and the legitimate means to attain those goals. Individuals would strive to attain valued goals yet would not be able to do so legitimately because of structural or individual limitations, leading to deviant behaviours	The school ensures that its priorities and activities do not inadvertently encourage criminal or ASBs. This could pertain to any aspect of the school's learning or pastoral care or physical environment. The school discourages unhealthy competitiveness (preventing bullying), ensures that it recognises broad achievement (preventing disengagement, truancy and substance use), ensures a safe environment (preventing violence) and socially supports students preventing (anxiety and substance use, etc.)	Antisocial outcomes (under-16 sex or smoking, under-18 drinking, illicit drug use, violence) and emotional/mental health outcomes	Sources of strain are societal. But can ameliorative action be taken at the level of the school without addressing societal structures?
Theory of reasoned action Ajzen and Fishbein 1980 ¹¹²	Behaviour is largely determined by behavioural intentions, which in turn depend on a person's attitudes about the behaviour (beliefs about the consequences of performing the behaviour along with valuation of these consequences) and subjective norms (perceived expectations from relevant individuals or groups along with intentions to comply with these expectations)	Health education teaches students about the social and health consequences of a wide range of healthy and risky behaviours, and encourages students to value health and reduce risk. School punishes some unhealthy behaviours. School communicates norms about healthy and risky behaviours. School ensures high student engagement and connection with school so that students seek to comply with school norms	Any outcome	But do individuals rationally weigh up multiple factors before making a decision? Aren't they also unconsciously influenced by peers or habit?

Appendix 5 Data extraction and quality appraisal tools

Research question 1: theory studies

TABLE 15 Data extraction and quality appraisal tool for theory studies stage 2

Section A: Descriptive	
What is its name?	Add text
Who originated it?	Add text
What was its year of origin?	Add text
What constructs and pathways does it involve?	Add text
What were its disciplinary origins?	Add text
Is it linked explicitly to other higher-order or lower-order theories?	Yes (details)
	No (details)
Section B: Evaluative	
Are its constructs well specified or are they very general and/or very vague?	Yes (details)
	No (details)
Does it specify causal relations between its constructs?	Yes (details)
	No (details)
Is it simple – does it use a reasonably small number of components and inter-relations between them?	Yes (details)
	No (details)
Can it be used to define what <i>specific</i> aspects of the school institution are important in determining health?	Yes (details)
	No (details)
Is it comprehensive in terms of tracing the pathways through which the school as an institution affects health?	Yes (details)
	No (details)
Is it useful to understanding multiple health domains or is it specific to one or two domains?	Yes (details)
	No (details)
What are its assumptions?	Add text
Are these assumptions explicit or implicit?	Yes (details)
	No (details)

Research question 2: outcome evaluation studies

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2

Section A: Study details	
Study funding	Stated (write in to describe) Not stated
Study design	Cluster RCT Matched cluster comparison (cease data extraction) Process evaluation (cease data extraction and refer to University of East London) Other (describe) (cease data extraction)
Author and year	Enter details (include author and year as 'Smith 2000')
Data set details	Enter name [write in data set/project name (e.g. NELS)]
Section B: Description of intervention	
Does the intervention have any community and curriculum component?	Yes (cease data extraction and refer to Cochrane team) No
Name of intervention	Stated (write in name) Not stated
Intervention description	Stated (write in to describe) Not stated
Aim(s)/objective(s) of the intervention	Stated (describe) (write in, as stated by the authors) Not stated
Country in which intervention was implemented (select all that are applicable)	USA UK Australia Canada Norway China New Zealand South Africa Thailand Finland Taiwan Netherlands Hong Kong Israel Germany Other (specify)
Who does the intervention target to provide benefit? (This may differ from 'which group of individuals is targeted for recruitment to be research participants?')	Teachers only (cease data extraction) Students only (write in to describe year group, gender, etc.)

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

	Teachers and students
	Not stated
Is a named theory used to inform the intervention?	Yes (describe) [write in as stated by the author(s)]
	No
	Not stated
Intervention start date	Stated (write in to describe)
	Not stated
Intervention length	Stated (write in to describe)
	Not stated
Details of intervention providers external to the school (who, how many, how trained, etc.)	Stated (write in to describe)
	Not stated
Details of intervention providers internal to the school (who, how many, how trained, etc.)	Stated (write in to describe)
	Not stated
Intervention costs	Stated (write in to describe)
	Not stated
What do comparison groups receive?	Stated (write in to describe)
	Not stated
Type of school	Stated (write in to describe)
	Not stated
Age range school caters for	Stated (write in to describe)
	Not stated
Section C: Study recruitment	
Sample size calculation	Stated
	Not stated
Recruitment of schools to study	Random sampling (write in to describe)
	Convenience sampling (write in to describe)
	Purposive or quota sampling (write in to describe)
	Other (write in to describe)
	Not stated
Incentives to schools	Yes (write in to describe)
	No
	Not stated
School response rate for recruitment into study	Stated (write in to describe, giving n/N and % if possible)
	Not stated
Which group of individuals is targeted for recruitment to be research participants? (This may differ from 'who does the intervention target to provide benefit?')	Teachers only
	Students only
	Teachers and students (write in to describe)
	Not stated
	<i>continued</i>

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

Sampling of evaluation study sample	Purposive or quote sampling (e.g. year groups) (write in to describe)
	Random sampling (write in to describe)
	Other (write in to describe)
	Not stated
Any exclusions regarding recruitment of individuals within schools to study sample, e.g. literacy, children on detention	Yes (write in to describe)
	No
	Not stated
Incentives to individual research participants	Yes (write in to describe)
	No
	Not stated
Informed consent for data collection from individual research participants	Yes (write in to describe)
	No
	Not stated
Individual research participants' response rate for baseline data	Stated overall (write in to describe, giving <i>n/N</i> and % if possible)
	Not stated overall
	Stated by arm (write in to describe, giving <i>n/N</i> and % if possible)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe)
	No
Number of schools per arm	Stated
	Not stated
Number of students per arm	Stated
	Not stated
Section D: Study sample baseline description overall and by arm	
School characteristics	Stated
	Not stated
Age	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
Gender	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Socioeconomic status	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Ethnicity	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Family	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Geographic area	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Baseline health status	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>

continued

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (continued)

	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Baseline education	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Other	Stated overall (write in to describe)
	Not stated overall
	Stated by arm (write in to describe)
	Not stated by arm
	<i>Significant baseline differences?</i>
	Yes (write in to describe, giving <i>p</i> -values if possible)
	No
Section E: Random allocation	
If random how was random sequence generated?	Computer (write in to describe)
	Random number tables (write in to describe)
	Other (write in to describe)
	Not stated
	Not applicable
What was the form of randomisation?	Not reported
	Matched
	Stratified
	Other
	Not applicable
Random allocation carried out before or after baseline survey?	Before
	After
	Not stated
Was the allocator blind to random allocation? Could they determine which school is allocated to which arm in advance?	Yes
	No
	Not clear
	Not stated
Were school participants blind to allocation? Did they know to which arm they were allocated?	Yes
	No
	Not clear
	Not stated

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

Section F: Follow-up data	
How many waves of follow-up occurred post baseline?	Stated (write in to describe)
What was the response rate for schools for each wave? Did any schools drop out?	Stated overall
	Not stated overall
	Stated by arm
	Not stated by arm
What was the response rate for individual research participants?	Stated overall
	Not stated overall
	Stated by arm
	Not stated by arm
Section G: Health outcome measurement	
Analytical status of measure	Primary indicated (please state) (write in to describe authors' interpretation, not reviewer's)
	Primary not indicated
	Secondary indicated (please state) (write in to describe – authors' interpretation, not reviewer's)
	Secondary not indicated
	Other (please describe) (write in to describe)
Health outcome measure (select as many as applicable)	Obesity, bodyweight or BMI (write in to describe)
	Physical activity (write in to describe)
	Eating or drinking (non-alcoholic) (write in to describe)
	Emotional, psychological or mental health (write in to describe)
	Violence, aggression, bullying or harassment (write in to describe)
	Substance use smoking, alcohol, drugs (write in to describe)
	Sexual health or teen pregnancy (write in to describe)
	Health or well-being in general (write in to describe)
	Sun protection (write in to describe)
	Injury (write in to describe)
Category of outcome	Other (write in to describe)
	Mental or physical health end point (write in to describe)
	Health-related behaviour (write in to describe)
	Health-related knowledge, attitude, perception of peer norms or other cognition (write in to describe)
Measurement	Other (write in to describe)
	Observation, physical examination or physical measurement (write in to describe)
	Researcher interview (write in to describe)
	Self-reported survey or diary (write in to describe)
<i>continued</i>	

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

	Practical or psychological test (write in to describe)
	Other (write in to describe)
	Not stated
Whether or not instrument validated/piloted	Yes, validated (write in to describe)
	No, not validated (write in to describe)
Was outcome measurement carried out blind to allocation?	Yes (write in to describe)
	No (write in to describe)
	Not stated
Whether or not measurement confidential	Yes (write in to describe)
	No (write in to describe)
	Not sure/clear
How many post-baseline measurements were taken?	Stated (write in to describe number and timing and which if any were being used as main study outcome)
	Not stated
Section H: Non-health outcome measurement	
What topic	Education attainment (write in to describe)
	Education-related attitudes or behaviour (write in to describe)
	Other
Measurement	Observation, physical examination or physical measurement (write in to describe)
	Researcher interview (write in to describe)
	Self-report survey or diary (write in to describe)
	Practical or psychological test (write in to describe)
	Other (write in to describe)
	Not stated
Whether or not instrument validated/piloted	Yes (write in to describe)
	No (write in to describe)
	Not stated
Was outcome measurement carried out blind to allocation?	Yes (write in to describe)
	No (write in to describe)
	Not stated (write in to describe)
Whether or not measurement confidential	Yes (write in to describe)
	No (write in to describe)
	Not clear (write in to describe)
How many post-baseline measures were taken?	Stated (write in to describe number and timing and which if any were being used as main study outcome)
	Not stated

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (continued)

Section I: Analysis	
Is a named theory used to inform study hypothesis/ research questions?	Yes (write in to describe) No Not stated
Do authors state in advance which outcomes they were intending to measure?	Yes [write in to describe, summarising which are primary, secondary, etc. (if relevant)] No
Do authors present outcome findings for all of the above outcomes?	Not relevant (do not state outcomes to be examined in advance) Yes, all (write in to describe) Yes, some (write in to describe) No (write in to describe) Other (write in to describe) Not clear
Level of analysis	Individual not adjusting for clustering (write in to describe) Individual adjusting for clustering (write in to describe) Multilevel (write in to describe) School level (write in to describe) Other (write in to describe) Not stated
Scope of analysis	Intention to treat On treatment Other (write in to describe) Not clear
Was outcome analysis carried out blind to allocation?	Yes (write in to describe) No (write in to describe) Not clear
Were there any significant differences in baseline measures?	Yes, noted above Yes, not noted above No Not stated
If there were significant baseline differences, were these adjusted for?	Not appropriate, no significant differences Yes, all (write in to describe what measures adjusted for and method of adjustment, e.g. logistic regression) Yes, some (write in to describe what measures adjusted for and method of adjustment e.g. logistic regression) No Not clear Not applicable

continued

TABLE 16 Data extraction and quality appraisal tool for outcome evaluation studies stage 2 (*continued*)

Did the authors report an intraclass correlation coefficient?	Yes (describe) No
Section J: Overall judgement	
Was confounding minimised? Study can 'pass' if there are not significant baseline differences or if any significant differences are adjusted for	Yes No Unclear
Was overall attrition <30% (to the time point at which outcome findings were reported)?	Yes No Not clear (describe if unclear)
Did attrition vary by <10% between study arms (to the time point at which outcome findings were reported)?	Yes No Unclear
Did authors report data on all outcomes they were intending to measure?	Yes No Unclear
Whether or not there are any baseline differences	Yes (write in to describe) No (write in to describe) Not analysed Not stated
Risk of bias by outcome	Stated Not stated
Section K: Final comments	
Reviewer's comment(s). Add in here any comments you may have on issues not covered by the preceding questions, as well as your general impressions of the study	Comments (describe) No comments
Is the study key worded correctly? In the light of the data extracted please check whether or not any of the existing keywords are superfluous, or whether or not any additional keywords should be added. Need to ensure that the way in which the study was coded for the mapping was carried out accurately; if not, this should be amended	Yes No (insert changes)
Flag study as relevant for intervention process evaluation review if it contains or refers to process evaluation – pass on to University of East London	Yes, relevant No, not relevant

Research question 3: process evaluation studies

TABLE 17 Data extraction and quality appraisal tool for process evaluation studies stage 2

Section A: Administrative details	
Name of reviewer	Specify
Date of coding	Details
Form of report	Journal article Report Book (chapter or whole) Manuscript or thesis Conference proceeding Multiple categories (because of linked papers); please specify the references of the linked papers Not stated
Section B: Study aim(s), rationale and context	
What are the research questions/aims of the study that relate to feasibility and acceptability (process evaluation) of the intervention? Please extract these as they are worded in the paper. Aims and objectives of the intervention are already extracted in tool for intervention outcome evaluations	Stated Not stated
When did the evaluation take place in relation to the intervention? Tick as many as appropriate. Specify further where possible. Provide exact dates of process evaluation start and end date when possible	Not stated Unclear Afterwards (please specify) Concurrently For a limited period during the intervention (please specify)
Which aspects of the intervention were evaluated? Tick as many as appropriate. Specify further where possible	Consultation/collaboration/partnerships (specify) Implementation/delivery of the intervention (specify) Management and responsibility (specify) Quality of the programme materials (specify) Skills and training of the intervention providers (specify) Other (specify)
Which attributes of the intervention were evaluated?	Costs associated with the intervention (specify) Perceptions, understanding and/or acceptability of the intervention (specify) Content of the intervention (specify) Accessibility of the intervention reach (specify) Other (specify)
Section C: Methods – sampling strategy, recruitment and analysis	
Were the methods used to recruit participants in the process evaluation described?	Yes (please specify) No (please specify) Unclear (please specify)

continued

TABLE 17 Data extraction and quality appraisal tool for process evaluation studies stage 2 (*continued*)

What was the total number of participants in the process evaluation?	Sample size details (please specify) Not provided Unclear (please specify)
What quantitative methods were used to collect data on the processes involved? Tick as many as appropriate. Specify further where possible. Identify using the info box the corresponding attribute or aspect (process) that is addressed for each data collection method	No quantitative methods used Closed question questionnaires (specify) Diaries (specify) Structured interview (specify) Observations/videos/audio (specify) Routine data (specify) Unclear (specify) Other (specify)
What qualitative methods were used to collect data on the processes involved? Tick as many as appropriate. Specify further where possible. Identify using the info box the corresponding attribute or aspect (process) that is addressed for each data collection method	No qualitative methods were used Unclear (specify) Documentation (specify) Focus group (specify) Interview (specify) Observations/videos/audio (specify) Self-completion report or diary/questionnaire (specify) Routine data (specify) Open-ended questionnaire (specify) Other (specify)
Were any named validated instruments used to collect data for quantitative analysis? Identify for which aspect or attribute of the process evaluation this guide has been used to evaluate	Yes (please specify) No (please specify) Not applicable
Were any established/named guides or tools used to collect data for qualitative analysis? Identify for which aspect or attribute of the process evaluation this guide has been used to evaluate	Yes (please specify) No (please specify) Not applicable
Who were the data collected from? Tick as many as appropriate. Specify further where possible. Identify the type of data that was collected for each group identified. You can refer to the section on attributes and aspects of evaluation. Provide details as necessary	Not stated Unclear (specify) Intervention provider (specify) Teachers Students Parents Other school staff (specify) Other (specify)
Do the authors describe any ways that they increase the rigour of their data collection tools? For example, mention previous validation of tools, published version of tools, involvement of target population in development of tools, piloting of tools (when more than one tool was employed, please provide details for each)	Yes (please specify) No (please specify)

TABLE 17 Data extraction and quality appraisal tool for process evaluation studies stage 2 (*continued*)

Were the methods of data analysis for quantitative measures described?	Yes (please specify) No (please specify) Not clear
Were the methods of data analysis for qualitative measures described?	Yes (please specify) No (please specify) Not clear
Were steps taken to increase rigour/minimise bias and error in the sampling for the process evaluation? Consider whether (a) the sampling strategy was appropriate to the questions posed in the process evaluation (e.g. was the strategy well reasoned and justified?); (b) attempts were made to include all relevant stakeholders and/or obtain a diverse sample (think about who might have been excluded who may have had a different perspective to offer); (c) characteristics of the sample critical to the understanding of the study context and findings were presented, (i.e. do we know who the participants are in terms of, for example, role in the intervention/evaluation, basic sociodemographics, etc.)	Yes, a fairly thorough attempt was made (please specify) Yes, several steps were taken (please specify) Yes, a few steps were taken (please specify) Unclear (please specify) No, not at all/not stated/can't tell (please specify)
Were steps taken to increase rigour/minimise bias and error in the data collected for the process evaluation? Consider whether (a) data collection tools were piloted/(if quantitative) validated; (b) data collection was comprehensive, flexible and/or sensitive enough to provide a complete and/or vivid and rich description/evaluation of the processes involved in the intervention [e.g. Did the researchers spend sufficient time at the site/with participants? Did they keep 'following up'? Were steps taken to ensure that all participants were able and willing to contribute? (e.g. confidentiality, language barriers, power relations between adults and young people); Was more than one method of data collection used? Was there a balance between closed and open-ended data collection methods?]	Yes, a fairly thorough attempt was made (please specify) Yes, several steps were taken (please specify) Yes, a few steps were taken (please specify) Unclear (please specify) No, not at all/not stated/can't tell (please specify)
Were steps taken to increase rigour/minimise bias and error in the analysis of the process data? Consider whether (a) data analysis methods were systematic (e.g. was a method described/can a method be discerned?); (b) diversity in perspective was explored; (c) the analysis was balanced in the extent to which it was guided by preconceptions or by the data, i.e. participants' views, researcher observations, etc.; (d) the analysis sought to rule out alternative explanations for findings (in qualitative research this could be carried out by, for example, searching for negative cases/exceptions, feeding back preliminary results to participants, asking a colleague to review the data, or reflexivity; in quantitative research this could be carried out by, for example, significance testing)	Yes, a fairly thorough attempt was made (please specify) Yes, several steps were taken (please specify) Yes, a few steps were taken (please specify) Unclear (please specify) No, not at all/not stated/can't tell (please specify)

Section D: Findings

What are the findings of the process evaluation related to aspects? Provide details of findings in any/all sections that are addressed	None Unclear <i>Consultation/collaboration/partnerships</i> Not applicable Findings specified Findings not specified <i>Implementation/delivery of the intervention</i> Not applicable Findings specified
--	---

continued

TABLE 17 Data extraction and quality appraisal tool for process evaluation studies stage 2 (continued)

	Findings not specified
	<i>Management and responsibility</i>
	Not applicable
	Findings specified
	Findings not specified
	<i>Quality of the programme materials</i>
	Not applicable
	Findings specified
	Findings not specified
	<i>Skills and training of the intervention</i>
	Not applicable
	Findings specified
	Findings not specified
	<i>Other (specify)</i>
	Not applicable
	Findings specified
	Findings not specified
What are the findings of the process evaluation related to attributes? Provide details of findings in any/all sections that are addressed	<i>Perceptions, understanding and/or acceptability of the intervention</i>
	<i>Costs associated with the intervention</i>
	Not applicable
	Findings specified
	Findings not specified
	<i>Content of the intervention</i>
	Not applicable
	Findings specified
	Findings not specified
	<i>Accessibility of the intervention reach</i>
	Not applicable
	Findings specified
	Findings not specified
Were the findings of the process evaluation grounded in/supported by the data? Consider whether (a) enough data are presented to show how the authors arrived at their findings; (b) the data presented fit the interpretation/support claims about patterns in data; (c) the data presented illuminate/illustrate the findings; (d) (for qualitative studies) quotes are numbered or otherwise identified so that the reader can see that they do not just come from one or two people	Reasonably well grounded/supported (please specify)
	Fairly well grounded/supported (please specify)
	Limited grounding/support (please specify)
Please rate the findings of the process evaluation in terms of their breadth and depth (it may be helpful to consider 'breadth' as the extent of description and 'depth' as the extent to which data have been transformed/analysed). Consider whether (a) range of processes/issues were covered in the evaluation; (b) the perspectives of participants are fully explored in terms of breadth (contrast of two	Limited breadth and depth
	Good/fair breadth but very little depth
	Good/fair depth but very little breadth
	Good/fair breadth and depth

TABLE 17 Data extraction and quality appraisal tool for process evaluation studies stage 2 (*continued*)

or more perspectives) (and depth) insight into a single perspective; (c) both the strengths and weaknesses of the intervention are described/explored; (d) the context of the intervention has been fully described/explored; (e) richness and complexity have been portrayed (e.g. variation explained, meanings illuminated); (f) there has been theoretical/conceptual development	
To what extent does the process evaluation privilege the perspectives and experiences of young people? Consider whether (a) young people are included in the process evaluation; (b) there was a balance between open-ended and fixed-response options; (c) young people were involved in designing the research; (d) there was a balance between the use of an a priori coding framework and induction in the analysis; (e) the position of the researchers (did they consider it important to listen to the perspectives of young people?); (f) steps were taken to assure confidentiality and put young people at their ease	Not at all A little (please specify) Somewhat (please specify) A lot (please specify)
Overall, what weight would you assign to this process evaluation in terms of the reliability of its findings? Think mainly about the answers given to the previous three questions	Low Medium High
What weight would you assign to this process evaluation in terms of the usefulness of its findings? Think mainly about the answers given to the previous three questions and (a) how well intervention processes are described (e.g. does it provide useful information on barriers and facilitators to implementation – factors that others implementing the intervention would need to consider?); (b) whether the findings can help us to explain the relationship between intervention process and outcome (e.g. why the intervention worked or did not work; factors influencing effectiveness; how the intervention achieved its effects)	Low Medium High

Research question 4: multilevel studies

TABLE 18 Data extraction and quality appraisal tool for multilevel studies stage 2

Section A: Study details	
Author and year	Details
Dissertation	Yes No
Why do you think this is a multilevel model study (tick all that apply)	Reference to multilevel models (write in to describe) Reference to structural equation models (write in to describe) Reference to hierarchical linear models (write in to describe) Reference to nested regression models (write in to describe) Reference to random coefficient (write in to describe) Reference to random-effects models (write in to describe) Reference to HLM, MPlus or MLwiN software Other (write in to describe and consult with colleague)

continued

TABLE 18 Data extraction and quality appraisal tool for multilevel studies stage 2 (*continued*)

Age range for which schools cater	Stated (write in to describe – e.g. age 11–16; if it goes outside 4–18 consult with colleague) Not stated
Section B: Context	
Year of data collection	Stated (write in to describe – e.g. 2005, 2006–8) Not stated
Location/country, city/region, e.g. Victoria, Australia; Texas, USA	USA UK Australia Canada Norway China New Zealand South Africa Thailand Finland Taiwan Netherlands Hong Kong Israel Germany Other specify
Type of school(s)	Stated (write in to describe – e.g. public or private, selective or comprehensive, primary or secondary) Not stated
Data set [name the data set/project (e.g. NELS)]	Details
Section C: Theory	
Is a named theory referred to in the report? [write name and author(s)]	Yes (write in to describe) No
Is this theory described?	Yes (write in to describe) No Not applicable (no named theory)
Is this theory used to inform hypotheses?	Yes (write in to describe) No Not clear Not applicable
Is this theory used to inform measures?	Yes (write in to describe) No Not clear Not applicable

TABLE 18 Data extraction and quality appraisal tool for multilevel studies stage 2 (continued)

Section D: Methods	
Research hypotheses/aims/questions	Yes (write in to describe) No
Design regarding this multilevel model analysis	Cross-sectional Cohort/longitudinal including secondary analysis of trial data Other (write in to describe) Not stated
Is there a sample size calculation?	Yes (write in to describe) No
Sampling of schools	Random or probability sampling (write in to describe) Purposive or quota sampling (write in to describe) Other (write in to describe) Not reported
Are there inclusion criteria for the schools?	Yes (write in to describe) No Not clear
Are there exclusion criteria for the schools?	Yes (write in to describe) No Not clear/specified
Sampling of students	All students in school Purposive or quota (write in to describe, e.g. all students in particular year groups) Random or probability sampling (write in to describe) Other (write in to describe) Not stated/specified
Are there inclusion criteria for students?	Yes (write in to describe) No Not clear/specified
Are there exclusion criteria for students?	Yes (write in to describe) No Not clear/specified
Did students/parents give informed consent for data collection?	Yes (write in to describe) No Not mentioned
Section E: Measures	
What objective* school-level exposures are included in the multilevel model tick all that apply and write in to describe. (*Exposure that is measured using routine data, researcher assessment or some other assessment made other than through aggregation of questionnaire self-reports from school students)	School type (write in to describe) School leadership and management (write in to describe) School inspection rating (write in to describe)

continued

TABLE 18 Data extraction and quality appraisal tool for multilevel studies stage 2 (continued)

	Teaching policies and practices (write in to describe)
	Academic attainment (write in to describe)
	Pastoral care or social support policies or practices (write in to describe)
	Student break time/playtime policies or practices (write in to describe)
	Discipline policies or practice (write in to describe)
	Physical environment (write in to describe)
	Other (write in to describe – include 'school size')
	Truancy
	School size
Section F: Follow-up and response rates	
School response rate at baseline	Stated (write in to describe numerator and denominator and % if details are given)
	Not stated
Student response rate at baseline	Stated (write in to describe numerator and denominator and % if details are given)
	Not stated
Average students per school in the multilevel model	Stated
	Not stated
Waves of follow-up post baseline	Stated (write in to describe numerator and denominator and % if details are given)
	Not stated
School response rate at each follow-up	Stated (write in to describe numerator and denominator and % if details are given)
	Not stated
Student response rate at each follow-up	Stated (write in to describe numerator and denominator and % if details are given)
	Not stated
Section G: Sample overall baseline characteristics	
Age	Stated (write in to describe)
	Not stated
Gender	Stated (write in to describe)
	Not stated
Socioeconomic status	Stated (write in to describe)
	Not stated
Ethnicity	Stated (write in to describe)
	Not stated
Family	Stated (write in to describe)
	Not stated
Geographic area	Stated (write in to describe)
	Not stated

TABLE 18 Data extraction and quality appraisal tool for multilevel studies stage 2 (*continued*)

Baseline health status	Stated (write in to describe) Not stated
Baseline education	Stated (write in to describe) Not stated
Other	Stated (write in to describe) Not stated
Section H: Which of these levels are included in the multilevel model?	
Which of these levels are included in the multilevel model?	Student (numbers extracted above) Class (write in number of classes in the multilevel model) Teacher (write in number of teachers in the multilevel model) School (numbers extracted above) City (write in number of cities in the multilevel model) Region (write in number of regions in the multilevel model) Other (write in to describe; write in number of these units in the multilevel model)
Section I: Complete questions below for each health outcome that is used in the multilevel model	
Outcome	Sample mean of outcome ICC Appropriate covariates Inappropriate covariates Goodness of fit Between-school variance explained by school-level variables School-level determinant (for each) Effect type = effect size Measure of precision Author conclusions about the effects of objective school-level exposures on health Stated (write in) Not stated
Reviewer's comments	Yes, comments No, no comments

Research question 5: qualitative studies

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2

Section A: Administrative details	
Name of the reviewer	Details
Date of the coding	Details [provide details of date(s) data extraction took place]
Broad/specific focus. If the study has a broad focus and this data extraction focuses on just one component of the study, please specify this here	Not applicable (whole study is focus of data extraction) Specific focus of this data extraction (please specify)
Form of report	Journal article Report Book (chapter or whole) Manuscript or thesis Conference proceeding Multiple categories (because of linked papers); please specify which categories and ensure that the linked papers are captured
Section B: Study aims, rationale and context	
What are the study aims, research questions and/or hypotheses?	Primary research questions (please specify) Secondary research questions (please specify) Not stated/unclear (please specify)
Are student health outcome(s) examined in the report? Do the study aims or conclusions suggest that the study is focused on drawing conclusions or developing theory about how the school environment structures or influences student health outcomes in our priority areas? Additional guidance: focus on core business of schools (teaching, discipline, pastoral care including student participation) and exclude studies focused merely on catering, PE, etc. (unless they have substantial discussion on how these impact on core business)	Yes No (cease data extraction)
Why was this study carried out at this point in time, in those contexts and with those people or institutions?	Rationale for study (please specify) Not stated/unclear (please specify)
Was the study informed by, or linked to, an existing body of empirical and/or theoretical research? Please write in authors' description if there is one. Elaborate if necessary, but indicate which aspects are the reviewer's interpretations. If different empirical and theoretical research was drawn upon for different research questions, please specify this and provide details	Yes (please specify) No (please specify) Not stated/unclear (please specify)
Which of the following groups were consulted in working out the aims to be addressed in the study?	Students/young people (please specify) Teachers (please specify) Education practitioners (please specify) Parents (please specify) Community representatives (please specify) Other (please specify) Not stated/unclear (please specify)

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (*continued*)

Do authors report how the study was funded?	Yes (please specify) No (please specify) Not stated/unclear (please specify)
In which country or countries was the study carried out? Please select those that apply	USA UK Australia Canada Norway China New Zealand South Africa Thailand Finland Taiwan Netherlands Hong Kong Israel Germany Other (write in to describe)
Around which main health topic is the study framed by its authors?	Obesity, bodyweight or BMI (write in to describe) Physical activity (write in to describe) Eating or drinking non-alcohol (write in to describe) Emotional, psychological or mental health (write in to describe) Violence, aggression, bullying or harassment (write in to describe) Substance use smoking, alcohol or drugs (write in to describe) Sexual health or teen pregnancy (write in to describe) Health or well-being in general (write in to describe) Sun protection (write in to describe) Injury (write in to describe) Other (write in to describe)
Section C: Methods – sampling strategy, recruitment and consent	
What methods were used to sample schools? Specify frame, overall approach, recruitment methods	Details of school recruitment methods (please specify) Not stated/unclear (please specify) Not applicable (please specify)
What methods were used to sample individuals? Specify frame, overall approach, recruitment methods, etc., e.g. letters of invitation, telephone contact, face-to-face contact	Details of individual recruitment methods (please specify) Not stated/unclear (please specify)
continued	

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (*continued*)

	Not applicable (please specify)
Were any incentives offered to schools or individuals?	Details of incentives provided (please specify) Not stated/unclear (please specify) Not applicable (please specify)
Who, if any, gave consent for schools and/or individuals? Please comment on the quality of consent if relevant	Participant consent sought Parental consent sought Other consent sought Consent not sought Not stated/unclear (please specify) Not applicable (please specify)
Are there ethical concerns about the way that the study was carried out? Consider consent, funding, privacy, etc.	Yes (please specify) No concerns
Section D: Contextual information – actual sample	
When was the study carried out?	Details of study timing (please specify) Not stated/unclear (please specify)
Whose health and well-being is being considered in this study?	Teachers only (cease data extraction) Parents only (cease data extraction) Student only (please specify) Teachers and students (please specify) Students and parents (please specify) Other (please specify) Not stated
What school level is being examined in the context of student health?	Primary Middle Secondary Kindergarten Nursery Infants Elementary Junior Junior high High Senior high No level specified, but age/grade range indicated Other (please specify) Not stated/unclear please specify
What is the sample size? State for each group, e.g. 20 teachers interviewed and 50 students interviewed	Stated Not stated
What is the geographical location of the school/individuals in the sample?	Urban (please specify)

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (*continued*)

	Rural (please specify)
	Not stated/unclear (please specify)
What ages/grades are covered by the actual sample? If more than one group, please describe for each group. Ages and/or grades to be described (e.g. 10–15 year olds; grades 8–10)	Stated Not stated
What is the sex of participants? Please give the numbers of the sample that fall within each of the given categories. If necessary refer to a page number in the report (e.g. for a useful table). If more than one group is being compared, please describe for each group	Female only (please specify) Male only (please specify) Female and male (please specify) Not stated/unclear (please specify)
Is the socioeconomic status of the individuals within the actual sample stated? If more than one group is being compared, please describe for each group	Yes (please specify) No, but the socioeconomic details at the school/institution level are given (please specify) No details provided Unclear (please specify)
Is the ethnicity of the individuals within the actual sample stated? If more than one group is being compared, please describe for each group	Yes (please specify) No, but details of ethnicity are given at the school/institution level (please specify) No details provided Unclear (please specify)
Is the school size stated? Actual numbers of students, (e.g. $n=1000$) or 'large school', etc.	Stated (write in and describe) Not stated
Is there any other useful information about the study participants or the schools/institutions from which they were recruited? [e.g. area, health status, private vs public, religious (Catholic)]	Yes (please specify) No (please specify) Unclear (please specify)
Are there any other details relevant to recruitment and consent?	Yes (please specify) No
Section E: Methods – data collection	
Which methods were used to collect the data? Please indicate all that apply and give further details where possible	Focus group One-to-one interview – face to face One-to-one interview – telephone Observation Self-completion open-ended questionnaire Self-completion report or diary Draw and write Other (please specify)
Are details provided of data collection methods or tool(s)/ guides? Please provide details including names for all tools used to collect data, and examples of any questions/items given. Also, please state whether source is cited in the report	Yes (please specify) No (please specify) Unclear (please specify)
Who collected the data? Please indicate all that apply and give further detail where possible	Researcher (please specify) Head teacher/senior management (please specify)
	<i>continued</i>

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (continued)

	Teachers (please specify)
	Parents (please specify)
	Pupils/students (please specify)
	Other educational practitioner (please specify)
	Other (please specify)
	Not stated/unclear (please specify)
Do the authors describe any strategies they used to increase the rigour of their data collection? (e.g. negative case analysis, checking results with participants)	Yes (please specify)
	No (please specify)
	Not stated/unclear (please specify)
Where were the data collected?	School (please specify)
	Home (please specify)
	Community setting (please specify)
	Other (please specify)
	Not stated/unclear (please specify)
Any other important features of data collection? (e.g. use of video or audio tape; ethical issues such as confidentiality)	Yes (please specify)
	No (please specify)
Section F: Methods – data analysis	
Which methods were used to analyse the data? Please give details, e.g. for in-depth interviews how were the data handled? Reviewers may be able to discern analysis methods from the form of findings presented even if the authors do not explicitly state their methods	Explicitly stated (please specify)
	Implicitly stated (please specify)
	Not stated/unclear (please specify)
Do the authors describe any strategies they used to increase the rigour of their data analysis? (e.g. negative case analysis, checking results with participants)	Yes (please specify)
	No (please specify)
	Not stated/unclear (please specify)
Who carried out the data analysis? Please give details of the person/people carrying out data analysis. Please capture any information presented on reflexivity/researcher bias	Explicitly stated (please specify)
	Implicitly stated (please specify)
	Not stated/unclear (please specify)
Section G: Quality assessment	
Were steps taken to increase rigour in the sampling? Consider whether (a) the sampling strategy was appropriate to the questions posed in the study (e.g. was the strategy well reasoned and justified?); (b) attempts were made to obtain a diverse sample of the population in question (think about who might have been excluded who may have had a different perspective to offer); (c) characteristics of the sample critical to the understanding of the study context and findings were presented (i.e. do we know who the participants were in terms of, for example, basic sociodemographics, characteristics relevant to the context of the study)	Yes, a thorough attempt was made (please specify)
	Yes, several steps were taken (please specify)
	Yes, a few steps were taken (please specify)
	Unclear (please specify)
	No, not at all/not stated/can't tell (please specify)
Were steps taken to increase rigour in the data collected? Consider whether (a) data collection tools were piloted/(if quantitative) validated; (b) (if qualitative) data collection was comprehensive, flexible and/or sensitive enough to provide a complete and/or vivid and rich description of people's perspectives and experiences (e.g. did the researchers spend	Yes, a thorough attempt was made (please specify)
	Yes, several steps were taken (please specify)
	Yes, a few steps were taken (please specify)
	Unclear (please specify)

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (*continued*)

sufficient time at the site/with participants? did they keep 'following up'? was more than one method of data collection used?); (c) steps were taken to ensure that all participants were able and willing to contribute (e.g. processes for consent, language barriers, power relations between adults and children/young people)	No, not at all/not stated/can't tell (please specify)
Were steps taken to increase rigour in the analysis of the data? Consider whether (a) data analysis methods were systematic (e.g. was a method described/can a method be discerned?); (b) diversity in perspective was explored; (c) (if qualitative the analysis was balanced in the extent to which it was guided by preconceptions or by the data); (d) the analysis sought to rule out alternative explanations for findings (in qualitative research this could be carried out by, for example, searching for negative cases/exceptions, feeding back preliminary results to participants, asking a colleague to review the data, or reflexivity; in quantitative research this could be carried out by, for example, significance testing)	Yes, a thorough attempt was made (please specify)
	Yes, several steps were taken (please specify)
	Yes, a few steps were taken (please specify)
	Unclear (please specify)
Were the findings of the study grounded in/supported by the data? Consider whether (a) enough data are presented to show how the authors arrived at their findings; (b) the data presented fit the interpretation/support claims about patterns in data; (c) the data presented illuminate/illustrate the findings; (d) for (qualitative studies) quotes are numbered or otherwise identified and the reader can see that they do not just come from one or two people	No, not at all/not stated/can't tell (please specify)
	Findings are well grounded/supported
	Findings are fairly well grounded/supported
Please rate the findings of the study in terms of their breadth and depth (it may be helpful to consider 'breadth' as the extent of description and 'depth' as the extent to which data have been transformed/analysed). Consider whether (a) range of issues are covered; (b) the perspectives of participants are fully explored in terms of breadth (contrast of two or more perspectives) and depth (insight into a single perspective); (c) richness and complexity have been portrayed (e.g. variation explained, meanings illuminated); (d) there has been theoretical/conceptual development	Limited grounding/support of findings
	Limited breadth or depth
	Good/fair breadth but very little depth
	Good/fair depth but very little breadth
To what extent does the study privilege the perspectives and experiences of children and/or young people? Consider the following questions and make an overall judgement: (a) there was a balance between open-ended and fixed-response options; (b) whether children/young people were involved in designing the research; (c) there was a balance between the use of an a priori coding framework and induction in the analysis; (d) the position of the researchers (did they consider it important to listen to the perspectives of children?); (e) steps were taken to assure confidentiality and put children/young people at their ease	Good/fair breadth and depth
	Not at all
	A little
	Somewhat
Overall, what weight would you assign to this study in terms of the reliability/trustworthiness of its findings? Guidance: think (mainly) about the answers you have given to the first four questions above	A lot
	Low
	Medium
What weight would you assign to this study in terms of the usefulness of its findings for this review? Guidance: think (mainly) about the answers you have given above and consider (a) the match between the study aims and findings and the aims and purpose of the synthesis; (b) its conceptual depth/explanatory power	High
	Low
	Medium
	High

continued

TABLE 19 Data extraction and quality appraisal tool for qualitative studies stage 2 (*continued*)

Section H: Findings	
What are the findings, themes, metaphors and/or theories? Try to distinguish the findings that demonstrate the key mechanism of the schools' influence on individuals, (i.e. potential causal pathways leading to certain health outcomes)	Provide details
Are there any other relevant findings of the study not already extracted? May not be how the school environment structures/influences student health, but findings considered related and important	Yes (please specify) No
How many of the findings are relevant to this review? Consider to what extent the study findings fit into the priority areas for this review and respond to the research question: through what process might school-level factors influence the health of students, staff, parents or the local community?	One or two sentences One or two sections Many findings
Reviewer comments	Comments No comments

Appendix 6 Data extraction and quality appraisal tables

Research question 1: theory studies

TABLE 20 Data extraction of theory studies study context

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Bonding theory/social bonding theory	Lothian 2003, ⁸⁰ Rountree and Clayton 1999 ⁷³	Hirschi 1969, ¹⁷⁹ Gottfredson and Hirschi 1990 ¹⁰⁷	1969, 1990	All individuals have some inclination to commit criminal acts but it is social bonds that prevent this. 'According to <i>control</i> or bond theories, a person is free to commit delinquent acts because his ties to the conventional order have somehow been broken' (Hirschi, 1969, p. 3). Essentially, the absence of strong, meaningful social bonds can lead to delinquency The four elements of social bond theory include (1) <i>attachment</i> (affection for prosocial peers, parents, family, school, etc.), (2) <i>commitment</i> (value placed on prosocial relationships and institutions), (3) <i>involvement</i> (the degree of time spent being involved in these conventional activities; the more time involved in conventional activities, the less time for non-conventional activities such as antisocial/criminal behaviour), (4) <i>belief</i> (in conventional values, morals, rules, etc.) Bonding to school is positive but is often harder for students from lower-class backgrounds because schools traditionally have a middle-class culture	Criminology	Rooted in the general theory of crime but with more focus on the peer group. Later developed into social control theory
Contagion theory/model	Lothian 2003, ⁸⁰ Monshouwer <i>et al.</i> 2007 ⁶⁹	Jencks and Mayer 1990 ¹⁸¹	1990	Suggests that behaviours (e.g. crime, ASB, delinquency) are 'contagious' through social networks, i.e. peer influences. The theory focuses on the role of social interaction and influences. For example, students may learn certain problem behaviours (e.g. substance use) from their peers. These problem behaviours 'spread' through these networks and peer influences. When a	Social psychology	Relates to social learning theory

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Control theory/ social control theory	Grunseit <i>et al.</i> 2008, ⁹⁷ Johnson and Hoffman 2000, ⁷¹ Kim 2001, ⁸⁸ Lothian 2003, ⁸⁰ Rountree and Clayton 1999, ⁷³ Way 2003 ⁸²	Hirschi 1969 ¹⁷⁹ See also Akers and Sellers 2004 ¹⁸²	1969	certain threshold of contagion is passed uptake becomes endemic Social control theory proposes that people's relationships, commitments, values, norms and beliefs encourage them to develop self-control and not break the law. Thus, if moral codes are internalised and individuals are tied into, and have a stake in, their wider community, they will voluntarily limit their propensity to commit deviant acts. The theory seeks to understand the ways in which it is possible to reduce the likelihood of criminality developing in individuals. It does not consider motivational issues, simply stating that human beings may choose to engage in a wide range of activities, unless the range is limited by the processes of socialisation and social learning	Criminology	Origins in social bonding theory and ultimately in functionalist theories of crime
Defiance theory	Way 2003 ⁸²	Sherman 1993 ¹⁸³	1993	'Defiance can occur under four conditions, all of which are necessary. The offender defines a criminal sanction as unfair. The offender is poorly bonded to or alienated from the sanctioning agent or the community the agent represents. The offender defines the sanction as stigmatizing and rejecting a person, not a law-breaking act. The offender denies or refuses to acknowledge the shame the sanction has actually caused him to suffer' (p. 460) ¹⁸³	Criminology	
Deterrence theory	Way 2003 ⁸²	Gibbs 1975 ¹¹¹	1975	The theory posits that an individual will be deterred from committing a crime if the consequences of the crime (i.e. punishment) outweigh any potential	Criminology	Rational choice theory

continued

TABLE 20 Data extraction of theory studies study context (continued)

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Differential association theory	Kim 2001 ⁸⁸	Sutherland 1924 ¹⁷⁶	1924	benefits. This theory suggests that crime and deviance can be controlled by punishment as long as the punishment is certain, severe and rapid. According to the theory, punishment can deter misbehaviour either specifically through directly experiencing a sanction or generally through knowledge about the punishments that others have received and the fear of consequences for inappropriate behaviours	Criminology	Relates to social learning theory
Ecological systems theory	Comeau 2005, ⁸⁵ Haley 2007, ⁷⁷ Haug <i>et al.</i> 2010, ⁸⁷ Khoury-Kassabri <i>et al.</i> 2009, ⁷⁸ Ishibashi 2009, ⁸³ Kumar <i>et al.</i> 2008, ⁶⁸ Leatherdale <i>et al.</i> 2010, ⁹⁰ Lothian 2003, ⁸⁰ Marsiglia <i>et al.</i> 2002, ⁹⁶ Wiium and Wold 2009 ⁷⁵	Bronfenbrenner 1979 ¹²	1979	This theory suggests that relationships and the environment of an individual impact and explain their development. The theory uses 'layers' to explain an individual's development Microsystems: 'is a pattern of activities, roles and interpersonal relations experiences by the developing person in a given setting with particular physical and material characteristics' (p. 22) ¹²	Developmental psychology	

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories	
				<p>Mesosystems: 'comprises the interrelations among two or more settings in which the developing person actively participates (such as, for a child, the relations among home, school and neighbourhood peer group; for an adult, among family, work and social life)' (p. 25)¹²</p> <p>Exosystem: 'refers to one or more settings that do not involve the developing person as an active participant, but in which events occur that affect, or are affected by, what happens in the setting containing the developing person' (p. 25)¹²</p> <p>Macrosystem: 'refers to consistencies, in the form and context of lower-order systems (micro-, meso- and exo-) that exist, or could exist, at the level of subculture or the culture as a whole, along with any belief systems or ideology underlying such inconsistencies' (p. 26)¹²</p> <p>Bronfenbrenner¹² argues that an ecological perspective has three distinguishing features: '(1) the developing person is a growing dynamic entity that progressively moves into and restructures the milieu in which he/she resides; (2) the environment exerts its influence requiring a process of mutual accommodation – characterized by reciprocity; (3) the environment is not limited to a single immediate setting but it extends to incorporate interconnections between settings as well as external influences emanating from larger surroundings' (p. 21)</p>			

continued

TABLE 20 Data extraction of theory studies study context (continued)

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Ecological model of co-ordinated school health programmes	Lohrmann 2008 ⁵⁷	Lohrmann 2008 ⁵⁷	2008	Health education and other areas of curriculum teach the knowledge and skills that students need to engage in healthy behaviours. The wider school context constitutes an immediate environment that supports or impedes students' ability to engage in these behaviours. The psychosocial climate, safety facilities and transportation are key to determining student safety, security and success. The psychosocial environment can be considered in terms of expectations, school ownership and bonding, conduct and discipline, opportunity to learn, student self-management and classroom management	Public health	Informed by Bronfenbrenner's ¹² ecological systems theory and Bandura's ¹⁷⁸ social cognitive model
Human functioning and school organisation	Aveyard <i>et al.</i> 2004, ⁶⁴ Markham <i>et al.</i> 2008, ⁶⁵ Bisset <i>et al.</i> 2007, ⁶⁶ Tobler <i>et al.</i> 2011, ⁵³ Gordon and Turner 2001 ⁹⁵	Markham and Aveyard 2003 ⁶²	2003	This theory is based partly on Bernstein's influential model of cultural transmission (Bernstein, 1977). Bernstein maintains that all schools aim to facilitate two inter-related but separate types of learning, or, in his terms, all schools aim to relay two orders; the instructional and regulatory orders. The instructional order focuses on the relaying of knowledge and skills, and is concerned with pupils' ability to contribute to future production through work. The regulatory order focuses on the relaying of values and is concerned with the conduct, character, and manner of pupils while they are at school and after they have left. In addition to pupils' own agency, there are two other influences on pupils' responses to the instructional and regulatory orders. The first is the social background of the pupil, particularly social class. The second is the	Sociology of education	Educational transmission of class theory

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
				<p>culture of the school – the methods used to convey the instructional and regulatory orders. Daniels, Holst, Lunt, and Ulsoe-Johansen (1996) drew on Bernstein's model of cultural transmission to show that pupils recognise the particular expectations and values of their school, and that these differed from school to school. If pupils accept and meet the demands of both the instructional and regulatory orders, they are termed committed. Committed pupils have the greatest opportunity to use school to promote good human functioning and health. If pupils reject or cannot meet the demands of both the instructional and regulatory orders they are termed alienated. Alienated pupils experience restricted opportunities to develop good human functioning and hence health. Pupils from middle class backgrounds are more likely to be committed, while those from working class backgrounds are more likely to be alienated. However, schools can influence the proportions of committed and alienated pupils. If schools reduce barriers between the school and the communities it serves, between pupils and teachers, between pupil groups, and between subjects ("weakly classified"), and if they increase pupils' input and control over learning ("weakly framed"), proportionately more pupils would be committed rather than alienated, when pupils' social class background was accounted for' (p. 1768)⁶⁴</p>		

continued

TABLE 20 Data extraction of theory studies study context (continued)

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Integrated perspective on delinquent behaviour	Henry <i>et al.</i> 2009 ⁶⁷	Elliot <i>et al.</i> 1979 ¹⁰⁸	1979	The major variables in the model include two types of social bonds (integration and commitment), attenuation experiences (such as failure in conventional social settings) and delinquent learning and performance structures. Two dominant aetiological paths to delinquency are postulated by the paradigm. The first path is characterised by weak integration into and commitment to the social order and by involvement in delinquent peer groups. Characteristics of the second major path are initially strong bonds to the social order, subsequent attenuation, or weakening, of those bonds, and exposure and commitment to delinquent peer groups	Criminology	Aims to expand and synthesise strain, social learning and social control perspectives
Moral authority theory	Way 2003 ⁸²	Durkheim 1997 [1897] ¹⁷⁴	1897	'According to Durkheim (1925/61), one of the most important roles for the school is the socialisation of children. Schools should not only teach socially appropriate behaviour but must also inculcate a general respect and obligation towards social rules. He refers to the social authority of rules as moral authority. Students who have internalised moral authority will respect and even appreciate the structure that discipline provides. Because it is internalised, moral authority facilitates self-control. For Durkheim, a successful school is one that does not need to enforce rules and dole out heavy punishments because students do not challenge authority of the rules in the first place' (p. 34) ⁸²	Sociology	

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Primary socialisation theory	Henry <i>et al.</i> 2009 ⁶⁷	Oetting and Donnermeyer 1991 ¹⁸⁴	1991	The models posit that children learn patterns of behaviour from primary socialisation sources, including school, family, and peer groups. To the extent that children and adolescents are bonded or attached to prosocial primary socialisation sources, their involvement in deviant behaviour, including drug use and precocious alcohol use, is attenuated because they are motivated to conform to the norms, expectations and values of the prosocial sources. On the other hand, weak bonds to prosocial sources and strong bonds to antisocial sources free young people from adhering to conventional norms that discourage alcohol use, and affected youth are more likely to follow the norms, expectations and values of antisocial sources (e.g. friends who are not strongly bonded to school). In sum, students with strong school bonds are more likely to delay the onset of alcohol use and are less likely to escalate the use of alcohol	Developmental psychology	Related to the theoretical underpinnings of social control theory, differential association theory and social learning theory
Problem behaviour theory	Haley 2007, ⁷⁷ Johnson and Hoffman 2000, ⁷¹ Kim 1999 ⁸¹	Jessor and Jessor 1977 ¹⁰⁹	1977	This theoretical framework is based on three explanatory variables: (1) the perceived environment system: distal and proximal social influence factors; (i.e. family and peer orientation); (2) the personality system: this includes social cognitions, individual values, expectations, beliefs and attitudes; (3) the behaviour system: conventional and problematic behaviour structures and problematic behaviour structures that work against one another; for example, the model suggests that young people engage in particular behaviours, such as drug use or risky sexual	Psychology of substance use	

continued

TABLE 20 Data extraction of theory studies study context (continued)

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Social capital theory	Gladden 2005 ⁸⁶	Coleman 1988 ¹⁰²	1988	behaviour, to achieve a goal; this personal goal could be to cope with low-self esteem, unhappiness, boredom or a sense of isolation/loneliness Social capital is a variety of entities with two elements in common: some aspect of social structure, and facilitating actions of actors within the structure. Social capital facilitates individual or collective action, generated by networks of relationships, reciprocity, trust and social norms. Social capital is a neutral resource that facilitates any manner of action. Social capital is strongest when networks involved are enclosed, intergenerational and stable and share values and reciprocal obligation is the norm		
Social cognitive theory	Dzewaltowski et al. 2009, ⁶³ Gastic 2005, ⁷⁶ Haley 2007, ⁷⁷ Ishibashi 2009 ⁸³	Bandura 1989 ¹⁷⁸	1969	Social cognitive theory is a learning theory based on the ideas that people learn by watching what others do and that human thought processes are central to understanding personality. Although there is a fair amount of influence on development generated by learned behaviour displayed in the environment in which one grows up, the individual person (and therefore cognition) is just as important in determining moral development. People learn by observing others, with the environment, behaviour and cognition all chief factors in influencing development. These three factors are not static or independent; rather, they are all reciprocal. For example, each behaviour witnessed can change a	Social psychology	Builds on social learning theory

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Social development model/theory	Henry <i>et al.</i> 2009, ⁶⁷ Johnson and Hoffman 2000, ⁷¹ Kim 1999 ⁸¹	Hawkins and Weis 1985 ⁹⁸	1985	person's way of thinking (cognition). Similarly, the environment that one is raised in may influence later behaviours, just as a father's mindset (also cognition) will determine the environment in which his children are raised Young people learn both antisocial and prosocial patterns of behaviour from the environment. There are three types of process variables in the social development model: opportunities for involvement, skills and reinforcements. It is these process variables that 'determine whether or not a youth's participation [in some area of life] will contribute to the development of a bond of attachment and commitment to and belief in conventional society' (p. 73) ⁹⁸	Social psychology	Derived from social learning theory and social control theory
Social disorganisation/integration theory	Akiba 2010, ⁷⁰ Bradshaw <i>et al.</i> 2009, ⁸⁹ Grunseit <i>et al.</i> 2008, ⁹⁷ Kairouz and Adlaf 2003, ⁷⁹ Kim 2001 ⁸⁸	Shaw and McKay 1942, ¹⁸⁵ Durkheim 1997 [1897] ¹⁷⁴	1942, 1897	Disorganised communities are characterised by poverty, ethnic heterogeneity and residential mobility, which produces weakened social stability and social controls, leading to the inability of communities to solve problems, which in turn lead to crime	Criminology	Relates to moral authority theory and social integration theory
Social learning theory	Brotman 2009, ⁹⁴ Gastic 2005, ⁷⁶ Johnson and Hoffman 2000, ⁷¹ Kim 1999 ⁸¹	Bandura 1977, ¹⁸⁰ Rotter 1964, ¹⁸⁶	1977, 1964	People learn through observing and or imitating others, assessing whether behaviour is associated with desired outcomes or with social superiors/role models	Social psychology	Informed social cognitive theory
Strain theory	Johnson and Hoffman 2000 ⁷¹	Durkheim 1997 [1897], ¹⁷⁴ Merton 1938 ¹¹⁰	1897, 1938	Strain theory states that social structures within society may encourage individuals to commit crime. Strain may be either structural, whereby processes at the societal level filter down and affect how the individual perceives his or her needs,	Criminology	Relates to moral authority theory and to social integration theory

continued

TABLE 20 Data extraction of theory studies study context (continued)

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Theory of planned behaviour	Lothian 2003, ⁸⁰ Wium and Wold 2009, ⁷⁵ Wium and Wold 2006 ⁷⁴	Ajzen 1985, ¹¹³ see also Ajzen 2012 ¹⁸⁷	1985	(i.e. if particular social structures are inherently inadequate or there is inadequate regulation, this may change the individual's perceptions as to means and opportunities), or individual, relating to the frictions and pains experienced by an individual as he or she looks for ways to satisfy his or her needs (i.e. if the goals of a society become significant to an individual, actually achieving them may become more important than the means adopted)	Social psychology	Extension of theory of reasoned action
Theory of reasoned action	Lothian 2003 ⁸⁰	Ajzen and Fishbein 1980 ¹¹²	1980	The model suggests that 'human action is guided by three kinds of considerations: readily accessible beliefs about the likely outcomes of the behaviour and the evaluations of these outcomes behavioural beliefs; readily accessible beliefs about the normative expectations and actions of important referents and motivation to comply with these referents normative beliefs, and readily accessible beliefs about the presence of factors that may facilitate or impede performance of the behaviour and perceived powers of these factors control beliefs' (p. 448) ¹⁸⁷ Behaviour is largely determined by behavioural intentions, which in turn depend on a person's attitude about the behaviour (beliefs about the consequences of performing the behaviour along with valuation of these consequences) and subjective norms (perceived expectations from relevant individuals or groups along with intentions to comply with these expectations)	Social psychology	Informs later theory of planned behaviour

Theory	Reference cited in	Originated by	Year	Constructs and pathways	Disciplinary origins	Links with other higher-order/lower-order theories
Theory of structuration	Fletcher <i>et al.</i> 2009 ⁹¹⁻⁹³	Giddens 1984 ¹⁸⁸	1984	Social structure both constrains and enables individual agency, the intended and unintended consequences of which are the reproduction or transformation of social structure	Sociology	
Theory of triadic influence	Murnaghan <i>et al.</i> 2007 ⁸⁴	Flay 1999 ⁵⁶	1999	Health behaviours are influenced by factors from three domains: intrapersonal factors (social competence and sense of self), socioenvironmental factors (behaviours of others and bonding to others) and the broader cultural environment (information and opportunities about behaviours and culture/religion). Each of these streams has distal and proximal elements, moving from the social-personal nexus to expectancies and evaluations, to cognitions and affect, and with dynamic inter-relationships between the streams. Attitudes, socially normative beliefs and self-efficacy determine decisions/intentions and behaviour	Public health	The theory of triadic influence integrates many sociological and psychological theories
Victimisation theory	Akiba 2010 ⁷⁰	Mendelson 1956 ¹⁸⁹	1956	Victimisation theory states that people who perceive themselves as vulnerable in the community are more likely to fear crimes	Criminology	

TABLE 21 Quality appraisal of theory studies

Theory	Constructs well specified	Causal relations	Simple	Can it suggest which specific aspects of the school institution might influence health?	Applicable to multiple health domains?	Assumptions	Assumptions explicit or implicit?
Bonding theory/social bonding theory	Yes	Yes	Yes	No – other than vague suggestion that the institution may be difficult to bond to for some students	No – applies only to delinquency outcomes	That all individuals are motivated to commit crimes but that prosocial bonding inhibits this	Explicit
Contagion theory/model	Yes	No – processes not defined	Yes	No – other than how schools group students through admissions and streaming	Yes	That schools norms are generally prosocial and healthy	Implicit
Control theory/social control theory	Yes	Yes	Yes	Partly – the ability of schools to foster attachment and inculcation of self-control and prosocial values, but no underlying mechanisms offered	No – applies only to delinquency outcomes	That individuals will be influenced by peers to initiate behaviours	Explicit
Defiance theory	Yes	Yes	Yes	No, other than perceived legitimacy of the discipline system	No – applies only to delinquency outcomes	That all individuals are motivated to commit crimes but that they vary in self-control. That low self-control and tendency to criminality are the same	Implicit
Deterrence theory	Yes	Yes	Yes	Partly – only that the severity of punishments may be influential	No – applies only to delinquency outcomes	That school norms are generally prosocial and healthy	Implicit
Differential association theory	Yes	Yes	Yes	No, other than how schools group students through admissions and streaming	No – applies only to delinquency outcomes	That outright defiance rather than more passive apathy is most important in delinquency	Implicit
						That deterrence is more important than building positive commitment to prosocial behaviour	Implicit
						That individuals will be influenced by peers to initiate behaviours	Explicit

Theory	Constructs well specified	Causal relations	Simple	Can it suggest which specific aspects of the school institution might influence health?	Applicable to multiple health domains?	Assumptions	Assumptions explicit or implicit?
Ecological systems theory	No – a general framework rather than a specific theory	No – processes not explained	No – many constructs and inter-relations	No	Yes	That interaction between individual and environment is more influential than individual traits	Explicit
Ecological model of co-ordinated school health programmes	No	No	No	Partly – generally vague but points to the importance of offering opportunities for physical activity	Yes	Not clear	Implicit
Human functioning and school organisation	Yes	Yes	No – many constructs and inter-relations	Yes – learning, discipline and pastoral care	Yes	That equipping students with autonomy will lead them to choose healthy behaviours	Explicit
Integrated perspective on delinquent behaviour	Yes	Yes	Yes	Partly – only in so far as whether or not students experience failure at school, whether or not schools have the ability to foster attachment and inculcation of self-control and prosocial values	No – applies only to delinquency outcomes	That both motivation and social bonds are important in determining behaviour That school norms are generally prosocial and healthy	Explicit Implicit
Moral authority theory	Yes	Yes	Yes	Partly – only in terms of schools' ability to foster general respect for authority, but no details about how	No – applies only to delinquency outcomes	That respect for authority is always a good thing and that it is clear what this 'authority' is That school norms are generally prosocial and healthy	Implicit Implicit
Primary socialisation theory	Yes	Yes	Yes	No	No – applies only to delinquency outcomes	That all individuals are motivated to commit crimes but that some are inhibited from doing so	Explicit

continued

TABLE 21 Quality appraisal of theory studies (continued)

Theory	Constructs well specified	Causal relations	Simple	Can it suggest which specific aspects of the school institution might influence health?	Applicable to multiple health domains?	Assumptions	Assumptions explicit or implicit?
Problem behaviour theory	Yes	Yes	Yes	Partly – only in terms of whether or not students experience problems with conventional behaviour, but not how schools determine this	Yes	That school norms are generally prosocial and healthy	Implicit
Social capital theory	Yes	Yes	Yes	Yes	Yes	That social ties constitute a transferable source of action	Explicit
Social cognitive theory	Yes	Yes	Yes	No, other than teachers as role models in student social learning	No – only public behaviours	That cognitions mediate the effects of social interactions on learning	Explicit
Social development model/theory	Yes	Yes	Yes	Yes – whether or not schools provide opportunities for students to participate, develop skills to participate and gain recognition and reinforcement	No – applies only to delinquency outcomes	That students would regard staff as superiors or models	Implicit
Social disorganisation/integration theory	Yes	Yes	Yes	No	No – applies only to delinquency outcomes	That bonding leads to positive behaviour, but unclear about the underlying mechanism	Implicit
Social learning theory	Yes	Yes	Yes	No, other than teachers as role models in student social learning	No – only public behaviours	That school norms are generally prosocial and healthy	Implicit
Strain theory	Yes	Yes	Yes	Partly – only in terms of whether or not schools provide legitimate opportunities for students to achieve their goals	No – applies only to delinquency outcomes	That ethnic and other heterogeneity is a form of instability	Implicit
						That instability must cause disorganisation	Explicit
						That students would regard staff as superiors or models	Implicit
						That school norms are generally prosocial and healthy	Implicit

Theory	Constructs well specified	Causal relations	Simple	Can it suggest which specific aspects of the school institution might influence health?	Applicable to multiple health domains?	Assumptions	Assumptions explicit or implicit?
Theory of planned behaviour	Yes	Yes	Yes	Partly – only regarding whether or not schools have rules with clear punishments concerning health behaviours	Yes	That individuals rationally weigh up multiple factors before making a decision	Explicit
Theory of reasoned action	Yes	Yes	Yes	Partly – only regarding whether or not schools have rules with clear punishments concerning health behaviours	Yes	That individuals rationally weigh up multiple factors before making a decision	Explicit
Theory of structuration	No – very general in scope	No – very general in scope	Yes	No	Not clear	That structure and agency are mutually constituted, not ontologically separate	Explicit
Theory of triadic influence	Yes	Yes	No – many constructs and inter-relations	Partly – access to knowledge through health education, opportunities for engaging in risk and rules/norms against this, opportunities for bonding with staff and other students, development of social skills through health and general education	Yes	That there are cultural, social and biology/personality drivers of behaviour	Explicit
Victimisation theory	Yes	Yes	Yes	No	No – violence only		

Research question 2: outcome evaluation studies

TABLE 22 Data extraction of outcome evaluation studies (study context)

Study	Intervention name	Type of school	Location
Battistich <i>et al.</i> 1989, ⁵⁵ Battistich <i>et al.</i> 1990, ⁵⁴ Battistich 2003, ¹²³ Solomon <i>et al.</i> 1988 ⁵⁸	CDP	Elementary schools (grades K–6)	USA (northern California)
Bonell <i>et al.</i> 2010 ⁶⁰	HSE	Secondary schools (grades 7–11)	UK (Greater London and south-east England)
Dzewaltowski <i>et al.</i> 2009 ⁶³	HYP	Middle schools (no details of grades)	USA (no further details)
Flannery <i>et al.</i> 2003, ¹¹⁹ Krug <i>et al.</i> 1997 ¹²¹	PeaceBuilders	Elementary schools (grades K–5)	USA (Pima County, AZ)
Flay <i>et al.</i> 2004 ⁴⁴	AAYP – SCI	Elementary/middle schools (grades K–8 except one school that was K–6)	USA (Chicago, IL)
Fonagy <i>et al.</i> 2009 ¹²⁰	CAPSLE	Elementary schools (grades K–5)	USA (KS)
Ridgers <i>et al.</i> 2007, ¹²⁴ Ridgers <i>et al.</i> 2007, ¹²⁵ Ridgers <i>et al.</i> 2010 ¹²⁶	Playground physical structure intervention	Elementary schools (grades K–4)	UK (north-west of England)
Sallis <i>et al.</i> 2003 ¹²²	M-SPAN	Middle schools (no details of grades)	USA (San Diego County, CA)
Stratton 2000 ¹¹⁷	Playground surface marking intervention	Primary	North-west England
Stratton and Mullan 2005 ¹¹⁸	Playground surface marking intervention	Primary	North-east Wales and north-west England

a Studies are grouped by their corresponding intervention name when applicable in subsequent tables.

TABLE 23 Data extraction of outcome evaluation studies (intervention details)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
CDP ^{4,55,58,123}	<p>Intervention aim is to encourage prosocial behaviour by providing children with experiences that engender a sense of community and climate of mutual respect and concern in the classroom and school. This is accomplished by establishing a caring classroom environment in which children are enabled to learn about others' needs, feelings and perspectives; to collaborate with one another and engage in prosocial actions; to discuss and reflect on their social experiences as they relate to values of fairness, kindness and social responsibility; and to exercise autonomy and participate in decisions about their activities and their classrooms. Teachers are trained (curriculum materials, workshops and coaching) in:</p> <p>(1) Co-operative activities, in which small groups of children work together towards common goals on academic and non-academic tasks, and are explicitly encouraged to strive for fairness, consideration and social responsibility. Includes opportunities for prosocial action, including students helping maintain and improve the school environment and performing charitable community service activities. Children above first grade have been assigned differentiated roles (e.g. facilitator, reporter, recorder) when in groups larger than two, to help</p>	No	5 academic years	Not stated	Classroom teachers	Not stated

continued

TABLE 23 Data extraction of outcome evaluation studies (intervention details) (continued)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
	<p>orchestrate group interaction. Teachers are trained in group interaction skills and are provided with opportunities to reflect upon this in pre- and post-session discussions of group process</p> <p>(2) Developmental discipline, classroom management that builds positive interpersonal relationships, involves children in rule-setting and decision-making and uses non-punitive control techniques that employ induction, mutual problem-solving and use of the minimum pressure necessary to gain compliance</p> <p>(3) Activities promoting social understanding, in which class meetings, discussion of books and films, and events that arise spontaneously in class are used to enhance sensitivity to, and understanding of, the feelings, needs and perspectives of others</p> <p>(4) Highlighting prosocial values and their expression in everyday life by pointing out exemplary behaviour in the classroom and in literature, films and television</p> <p>(5) Helping activities, in which children are encouraged to help others by carrying out classroom chores, helping other students in class, participating in peer tutoring and 'buddies' programmes, and performing</p>					

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
HSE ⁶⁰	charitable community activities and helping activities in the school at large The HSE intervention involves each intervention school instituting an action team to plan and deliver actions to improve students' relationships with teachers and other students, security, social support, self-regard and engagement, so as to reduce substance use. The action team comprised staff, students from various year groups and parents. The various actions were either mandatory (e.g. revising school policies and rules) or locally determined (e.g. developing 'safe spaces' for younger students, training older students as peer mediators, resolving disputes and preventing bullying). The intervention included a facilitator with accompanying manual, staff training, funding and a survey of student needs. There was no curriculum component	No	1 school year	Facilitator with head teacher experience; research staff to conduct and analyse needs survey	Action team (staff, students)	Continued with normal practice
HYP ⁶³	The HYP intervention aimed to develop the 'personal and proxy agency' of adult leaders and youth to build middle school environments (healthy places) that promote fruit and vegetable consumption and physical activity. This was to occur through building youth self-efficacy and youths' confidence so that they could influence others, teachers and parents, to assist them in building healthy places. Expert staff delivered continuous group staff training to paid school site co-ordinators from the	Yes – social cognitive theory	2 academic years or 3?	'Expert staff'	School staff and students	Not stated

continued

TABLE 23 Data extraction of outcome evaluation studies (intervention details) (continued)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
PeaceBuilders ^{119,121}	<p>eight intervention schools. School site co-ordinators were linked together as part of a 'performance community hub' to facilitate problem-solving with the help of four training sessions yearly and monthly conference calls. Site co-ordinators targeted a place for environmental change (school lunch, after-school programme, classroom) that would either build youths' environmental change skills or provide options for physical activity or fruit and vegetable consumption in a positive social environment. Sites promoted and marketed their environmental changes. A positive social environment was characterised by CASH (Connection, Autonomy, Skill building and Healthy norms). Site co-ordinators facilitated youth-led school advocacy groups, known as 'change teams', which were the hub of intervention activities at the school. Students implemented their environmental change efforts. Key adult place leaders also participated on the change teams. The change teams created awareness and visibility within their school regarding the importance of physical activity and good nutrition, developing site-specific promotional videos. A seventh- and eighth-grade curriculum (Students Building and Promoting Healthy Places) was implemented to help facilitate student leadership</p> <p>PeaceBuilders is a universal, elementary school-based violence prevention programme that attempts</p>	No	1 academic year (within the trial)	'External person ("model developer")'	School staff	Normal practice (waiting list control) with \$1000 as an incentive for them not to

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
AAYP – SCI ⁴⁴	<p>to alter the climate of a school by teaching students and staff simple rules and activities aimed at improving child social competence and reducing aggressive behaviour. PeaceBuilders focuses on individual behaviour change in proximal interpersonal and social settings. The intervention is woven into the school's everyday routine rather than presented as a time- or subject-limited curriculum. PeaceBuilders specifically rewards prosocial behaviours and provides strategies to avoid the differential or accidental reinforcement of negative behaviours. All children and staff in the school learn five simple rules using a common language, which makes the intervention very easy to learn and maintain: (1) praise people, (2) avoid put-downs, (3) seek wise people as advisers and friends, (4) notice and correct hurts we cause, and (5) right wrongs</p> <p>The SCI included a curriculum as per the SDC arm plus parental support, school climate and community components to impact all social domains of influence on children. Each SCI school formed a school task force consisting of school personnel, students, parents, community advocates and project staff to implement the programme components, propose changes in school policy, develop other school-community liaisons supportive of school-based efforts and solicit community organisations to conduct activities to support the SCI efforts. A</p>	Yes – theory of triadic influence	3 years	Not stated	School personnel, students, parents, community advocates and project staff	engage in any PeaceBuilders programme-related activities

continued

TABLE 23 Data extraction of outcome evaluation studies (intervention details) (continued)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
CAPSLE ¹²⁰	<p>goal of these linkages was to 'rebuild the village' and create a 'sense of ownership' by all stakeholders to promote sustainability of these efforts on completion of the project. Teacher and staff in-service training enabled them to review and model curriculum skills; integrate prosocial skills into the school environment; consider examples of school activities to reinforce curriculum skills; model proactive classroom management skills; promote interactive and cultural teaching methods; and promote a sense of self and purpose. The school task force implementation included reviews of school policy; school-wide fairs; annual field trips for programme parents and children; grant writing for local monies; and soliciting monies and supplies from local businesses. The parent training enabled them to reinforce skills taught in the SDC; improve child supervision and methods of discipline; enhance anger and stress management; enhance parent-child communication; and promote parent-teacher communication</p> <p>The CAPSLE intervention aims to improve the capacity of all community members to mentalise (interpret both one's own and others' behaviour in terms of beliefs, wishes, feelings) to counteract bullying of others. This is a manualised, psychodynamic social systems approach addressing the co-created relationship between bully, victim and bystanders, and assumes</p>	No	2 years of active intervention followed by a limited third year of intervention designed to mimic implementation outside of a research protocol	Study investigator; a team drawn from staff in schools were involved in the pretrial pilot	All school staff	Treatment as usual

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
Playground physical structure intervention ¹²⁴⁻¹²⁶	<p>that all members of the school community, including teachers, play a role in bullying. In year 1, teachers received group training, students received self-defence training and the CAPSLE team consulted with school staff monthly. Year 2 began with school-wide refresher training for all school staff and a refresher self-defence course, and ongoing consultation with counsellors and teachers and the adult/peer mentor programmes. There were biweekly supervision meetings with the intervention team during years 1 and 2. At the beginning of year 3 the CAPSLE team ceased working with the schools and only in-service refresher training and self-defence was provided</p> <p>Each school received £20,000 (total £300,000) to redesign the playground environment in to three specific zonal coloured areas: (1) a red zone, (2) a blue zone and (3) a yellow zone. The red zone was a designated sports area with physical markings for various sports, i.e. soccer, tennis, basketball. The blue zone served as a fitness and skills (multiactivity) area and the yellow zone served as a 'chill-out'/quiet play area. The intervention schools received the following physical structures: soccer goal posts, basketball hoops, fencing around the red sports area and seating in the yellow quiet area. None of the supervisors received training in the promotion of playground physical activity</p>	No	1 year	Not stated who redesigned and undertook playground zonal design although most likely external to the schools	School teachers (supervising morning and afternoon recess)	Comparison groups each received small items of sporting equipment (soccer balls, jump rope and tennis balls)

continued

TABLE 23 Data extraction of outcome evaluation studies (intervention details) (continued)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
M-SPAN ¹²²	This intervention aimed to (1) increase the total energy expenditure from physical activity by the student population at school and (2) decrease the grams of total and saturated dietary fat purchased at or brought to school by students. Intervention strategies were designed to change school policies and environments to provide more healthful food choices and more opportunities for physical activity at school. Because the study was designed to test environmental interventions, there was no classroom education. Key school personnel met with project staff to select and implement policy changes to create healthier school environments. Participants included administrators, (principals, food service directors), faculty (physical educators), staff (cafeteria managers, student body organisation advisors), parents and students. The project requested three, 90-minute meetings per school across 2 years. In each intervention year participants selected two to four policies to improve and formed action plans to achieve their goals. A manual guided the work of the groups. Example policies included 'Serve 1% or skim milk only', 'Close school stores at lunchtime', 'Provide supervision and transportation for student physical activity after school' and 'Upgrade PE facilities and equipment'. The schools' goals were printed on a poster-style newsletter that was distributed to staff at all intervention schools. Progress on	Yes – structural ecological model	2 years	Not stated	School staff	Not stated

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
	<p>goals was monitored in subsequent meetings. Student health committees were also formed, which consisted of 9–12 students and were supervised by a faculty member and project staff. Members received t-shirts and training booklets describing how committees could support healthy policies and promote healthful choices. The goal was to have a monthly activity, such as assisting with taste tests, announcing after-school activities and creating posters promoting healthful lunch options. Parental education was delivered through existing school communication channels and was conceptualised as changing the information environment. All 24 schools received an incentive to participate (\$1000 for PE equipment), and intervention schools received an additional \$500 for kitchen equipment and \$2000 for physical activity programmes or equipment. Outputs of policy changes included strategies to increase physical activity on campus during leisure periods throughout the school day, when students could make choices (i.e. before school, after lunch and after school) providing supervision, equipment and organised activities. In addition, PE was required daily in all grades, and interventions with school food service staff and managers were undertaken to provide more low-fat choices at these sources</p>					

continued

TABLE 23 Data extraction of outcome evaluation studies (intervention details) (continued)

Intervention	Intervention description	Theory used	Length	Providers external to the school	Providers internal to the school	Comparison groups
Playground surface marking ¹¹⁷	Children designed a series of markings that were painted in bright fluorescent colours on the tarmac playground surface. Ten markings were linked to school curriculum themes (castles, dragons, etc.) and were evenly spaced throughout the playground area	None	One-off design intervention	Magical Markings, Castleford, UK	None	Nothing
Playground surface marking ¹¹⁸	Playgrounds were painted during the summer holiday in bright fluorescent colours that varied according to school preference (castles, dragons, etc. in early primary 4–7 years; markings for sports in late primary 7–11 years). The cost was £800 per school	None	One-off design intervention	Grant from Health Promotion Wales	None	Nothing

TABLE 24 Data extraction of outcome evaluation studies (intervention data collection methods)

Intervention	Waves of data collection	Number of schools per arm at each wave	Number of students per arm at each wave
CDP ^{54,55,58,123}	Seven grades K–6 (no pre-intervention baseline), but not all reported on	Grade 3: intervention=3, comparison=3; grade 4: intervention=3, comparison=3; grade 5: intervention=1, comparison=1; grade 6: intervention=2, comparison=2	Grade 3: intervention=285, comparison=236; grade 4 (self-esteem): intervention=148, comparison=140; grade 6 (loneliness/social dissatisfaction and social anxiety): intervention=148, comparison=101 States that n=133 (39% of original sample) completed all follow-ups (analysis sample for Battistich <i>et al.</i> ⁵⁵)
HSE ⁶⁰	Two waves: baseline survey and 9-month follow-up	Year 7: intervention=2, comparison=2	Year 7: baseline=614 (295 intervention, 319 control); follow-up=735 (388 intervention, 347 control)
HYP ⁶³	Three	Eight per arm at each wave	Grade 6: intervention=1007 (74% of eligible), control=1204 (80%); grade 7: intervention=103 moved schools and 53 lost to follow-up, control=98 moved schools and 87 lost to follow-up; grade 8: intervention=60 moved schools and 24 lost to follow-up, control=70 moved schools and 134 lost to follow-up; intervention=767 (76%), control=815 (68%) For MVPA outcome, complete follow-up data on intervention=648, control=716; for fruit and vegetable consumption outcome, complete follow-up data on intervention=617, control=715 Baseline: intervention=1631 (90.7%), control=1105 (86.0%); follow-up: intervention=1392 (84.4%), control=1348 (77.9%) The average rate of attrition (autumn data) but no spring data was 12% in year 1. Rates of attrition from baseline were not significantly different by arm. New kindergarten students were not included in the follow-up sample and those in fifth grade at baseline were not followed up. For all students assessed at baseline, 169 (10.5%) of 1615 students (grades K–2) and 120 (10.5%) of 1140 students (grades 3–5) had no other data over the 2-year period
PeaceBuilders ^{119,121}	Two pertaining to RCT analysis	Intervention arm=4, control arm=5 (but one school with grades K–2 and nearby one with grades 3–5 formed single analytical unit) Additionally, a comparison school dropped out when allocated to control and was replaced with another (non-randomised) school with no baseline data	Not applicable for Krug <i>et al.</i> ¹²¹ analysis as draws on routine data from nurses' logs
AAYP – SC ¹⁴⁴	Five	Six per arm at each wave	Not stated Average sample of 644 students range (597–674) at each wave

continued

TABLE 24 Data extraction of outcome evaluation studies (intervention data collection methods) (continued)

Intervention	Waves of data collection	Number of schools per arm at each wave	Number of students per arm at each wave
CAPSLE ¹²⁰	Five	Three per arm at each wave	<p>Response rate 89.5–92.7% per wave</p> <p>339 (51%) of the 668 original grade 5 students still present at the end of grade 8</p> <p>Total analysis sample (students with one or more waves of data) of 1153</p> <p>Third to fifth graders</p> <p>CAPSLE: year 1 participating=391, lost (no data, no consent, absent)=87; follow-up from year 1 to year 2 (years 1–2) participating=196, lost=231; year 2 participating=356, lost=156; years 2–3 participating=216, lost=177; 3 participating=395, lost=135</p> <p>School Psychiatric Consultation active control: year 1 participating=296, lost=103; years 1–2 participating=145, lost=158; year 2 participating=283, lost=109; years 2–3 participating=184, lost=144; year 3 participating=323, lost=96</p> <p>Non-active control: year 1 participating=271, lost=71; years 1–2 participating=106, lost=156; year 2 participating=221, lost=115; year years 2–3 participating=119, lost=123; year 3 participating=217, lost=101</p>
Playground physical structure ^{124–126}	Four waves: baseline, 6 week post intervention, 6 months and 12 months	Ridgers <i>et al.</i> ^{124,125} report 15 intervention schools and 11 control schools at baseline, 6 weeks and 6 months; however, Ridgers <i>et al.</i> ¹²⁶ report 15 intervention schools at each follow-up but 10 control schools at 6 months' follow-up (with no explanation of disparities) as well as seven schools at 12 months' follow-up	<p>297 children wore accelerometers at baseline (intervention=149, non-intervention=148)</p> <p>Ridgers <i>et al.</i>¹²⁵ report that 242 children wore accelerometers at 6 weeks (intervention=106, non-interventions=136). Ridgers <i>et al.</i>⁹⁰ report that at 6 months the follow-up for heart rate and accelerometry was 84% and 92% for the intervention and control groups respectively. However, Ridgers <i>et al.</i>⁹² report 256 intervention and 270 control children at baseline; 205 intervention and 163 control children at 6 months' follow-up with no explanation of disparities with earlier reports; and 179 intervention and 88 control children at 12 months' follow-up</p>

Intervention	Waves of data collection	Number of schools per arm at each wave	Number of students per arm at each wave
M-SPAN ¹²²	Observational data – three; self-report data – two	12 per arm at each wave	Surveys were distributed to separate random samples at baseline (spring 1997) and 2 years later (spring 1999); age group unspecified Overall for self-report data, $n=1434$ student–parent pairs at baseline and 1434 at follow-up, not reported by arm
Playground surface marking ¹¹⁷	Pre-intervention baseline and post-intervention follow-up	One intervention school and one control school in each arm at each wave	Intervention arm: baseline=36, follow-up=27; control arm: baseline=24, follow-up=20
Playground surface marking ¹¹⁸	Pre-intervention baseline and post-intervention follow-up	Four intervention schools and four control schools in each arm at each wave	Intervention arm: baseline=120, follow-up=67; control arm: baseline=120, follow-up=32

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes)

Intervention	Study	Health outcome	Measurement	Follow-up
CDP	Battistich <i>et al.</i> 1989 ⁵⁵	General conflict resolution strategies (average combined score for method chosen and who this favoured across three situations)	All measures derived from private interviews. Interviewers and coders blind to allocation	Grades K–4
		% prosocial conflict resolution strategies (relative use of specific strategy)	Posed three hypothetical situations concerning potential conflict with another child to elicit strategies, probing for 'what if' should initial strategies not work. Taped and coded. Coded for (a) the strategy used (physical or verbal aggression, take object back, appeal to authority, do nothing, ask for return, share); (b) who was favoured (self, other, both) and (c) whose needs considered (own, others', both). 10% of codes checked, correlation > 75%	
		% antisocial conflict resolution strategies (relative use of specific strategy)		
		Consideration of others' needs in conflict resolution (average score for whose needs were considered with high scores indicating consideration of others)		
	Battistich 1990 ⁵⁴	Self esteem (new measure, alpha=0.87)	Questionnaire self-reports	Grade 4
		Loneliness/social dissatisfaction (established measure, alpha=0.90)		Grade 6
		Social anxiety (established measure, alpha=0.82)		Grade 6
	Battistich 2003 ¹²³	Loneliness and social dissatisfaction	Average response to a 16-item measure focused on relationships in school – higher score = greater feelings of loneliness and social dissatisfaction	Assessed at sixth grade
		Social anxiety (established measure)	Average response to a 10-item measure (Social Anxiety Scale for Children)	Assessed at sixth grade
		Self-esteem (established measure, internal consistency=0.87)	Average response to a 17-item measure was computed as an overall index – higher score = greater self-esteem	Assessed at fourth grade
Solomon <i>et al.</i> 1988 ⁵⁸	Students solve own minor interpersonal problems	Teacher questionnaire self-report	All grades combined	

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
Multivariate analysis of variance to control for multicollinearity in dependent measures	Not reported; differences suggested programme benefit	$F_{1,129} = 7.96, p < 0.006$	Information bias as (students not blind to allocation)	No baseline from student cohort so may be unmeasured confounding
	Not reported; differences suggested programme benefit	$F_{1,129} = 9.58, p < 0.003$	Information bias	No adjustment for cluster so likely to be overestimate of effects
	Not reported	Not significant	Information bias	$n =$ three schools per arm so potential for random error
	Not reported; differences suggested programme benefit	$F_{1,129} = 12.16, p < 0.0008$	Information bias	Attrition high so possible selection bias; not reported if differential by arm
Test for interaction of arm with grade to examine whether or not intervention effect increased with time	Not applicable; differences suggest that intervention benefits increased with grade	$F_{2,258} = 11.68, p < 0.0001$		
Multivariate analysis of variance	Not reported	Univariate $F = 1.47$, not significant	Information bias	See Battistich <i>et al.</i> ⁵⁵
	Not reported; differences suggested programme benefit	Univariate $F = 11.87, p < 0.01$		Only four out of six schools involved in grade 6 follow-up, with one intervention and one comparison school missing, so very low numbers
	Not reported; differences suggested programme benefit	Univariate $F = 4.74, p < 0.05$		
	Programme students less lonely and socially dissatisfied	$F_{1,233} = 11.58, p < 0.002$	Information bias as (students not blind to allocation)	See Battistich <i>et al.</i> ⁵⁵
	Programme students scored significantly lower than comparison students for social anxiety	$F_{1,233} = 4.92, p < 0.03$		
	Effect estimates and p -values not reported	Not significant		
No details	Not reported; differences suggested programme benefit	$T = 2.79, p < 0.001$	Information bias as teachers not blind to allocation	See Battistich <i>et al.</i> ⁵⁵

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
HSE	Bonell <i>et al.</i> 2010 ⁶⁰	Teased/threatened weekly or more and/or hurt ever in this school	Questionnaire self-reports (conducted during class time, in private) containing adapted items	Assessed at seventh grade and followed up approximately 9 months later
		Teased others in this school		
		Hurt others in this school		
		Feel safe in this school		
		Been in a fight at this school		
		Believes will try illegal drugs in the future		
		Believes will try smoking cigarette in the future		
		Believes will get drunk before age 16 years		
HYP	Dzewaltowski <i>et al.</i> 2009 ⁶³	% of 30-minute blocks of VPA (51 blocks across 3 days)	Validated Previous Day Physical Activity Recall (PDPAR) measure. PDPAR uses 30-minute blocks within a time grid to help children record physical activity over the previous day from 1500–2330 (17 blocks), rated by type of activity and intensity (light, moderate, hard, very hard)	Grades 6–8
		One or more VPA blocks per day		
		% of 30-minute blocks of MVPA (51 blocks across 3 days)		
		Two or more MVPA blocks per day standard		
		Fruit and vegetables, no. of servings per day	Validated Youth Adolescent Questionnaire assessing diet and heights and weight	
		Fruit, no. of servings per day		
		Vegetables, no. of servings per day		
		Students' proxy efficacy in getting parents to create supportive environments for physical activity, four items	Newly developed Youth Psychosocial Survey	

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
Intention-to-treat analyses adjusting for cluster reporting unadjusted and adjusted ORs, the latter adjusted for baseline sex, housing tenure, measures of each outcome	Adjusted OR=0.99	95% CI 0.59 to 1.51, not significant	Information bias (students not blind to allocation)	Selection bias – intervention and comparison school swapped as a result of former dropping out of intervention
	Adjusted OR=0.68	95% CI 0.46 to 1.01, $p=0.055$	Information bias (students not blind to allocation)	Random error as only two intervention and two comparison schools
	Adjusted OR=0.68	95% CI 0.46 to 1.04, $p=0.075$	Information bias intervention (students not blind to allocation)	
	Adjusted OR=2.89	95% CI 1.63 to 5.15, $p<0.001$	Information bias intervention (students not blind to allocation)	
	Adjusted OR=0.64	95% CI 0.38 to 1.09, not significant	Information bias intervention (students not blind to allocation)	
	Adjusted OR=0.71	95% CI 0.26 to 1.92, not significant	Information bias intervention (students not blind to allocation)	
	Adjusted OR=0.83	95% CI 0.43 to 1.61, not significant	Information bias intervention (students not blind to allocation)	
	Adjusted OR=0.86	95% CI 0.39 to 1.88, not significant	Information bias intervention (students not blind to allocation)	
Generalised linear mixed-model analysis. Condition (intervention vs control) and strata were modelled as fixed effects. School and time were included as random effects nested within strata. Comparisons of least-squares means were evaluated at $p<0.05$, two-tailed tests. Covariate terms were added to represent the main effects and their interaction with the strata and condition over time. Non-significant main effect covariates were included in the model. Non-significant interaction (four-way, three-way, two-way) terms were deleted in a stepwise deletion process. If the primary behavioural outcome was significant (MVPA, VPA, fruit and vegetable consumption), the extent to which self- and proxy efficacy mediated the impact of the intervention was assessed using a	Not reported; differences suggested programme benefit	Crude $p=0.11$; adjusted (sex, SES, ethnicity, BMI) $p=0.003$	Information bias as (students not blind to allocation)	Attrition overall is low to moderate but is slightly higher in control group
	Not reported	Crude $p=0.91$; adjusted (sex, SES, ethnicity, BMI) $p=0.31$	Information bias	
	Not reported; differences suggested programme benefit	Crude $p=0.49$; adjusted (sex, SES, ethnicity, BMI) $p=0.005$	Information bias	
	Not reported	Crude $p=0.99$; adjusted (sex, SES, ethnicity, BMI) $p=0.23$	Information bias	
	Not reported	Crude $p=0.75$; adjusted (sex, SES, ethnicity, BMI) $p=0.99$	Information bias	
	Not reported	Crude $p=0.97$; adjusted (sex, SES, ethnicity, BMI) $p=0.99$	Information bias	
	Not reported	Crude $p=0.88$; adjusted (sex, SES, ethnicity, BMI) $p=0.11$	Information bias	
Not reported	Crude $p=0.16$; adjusted (sex, SES, ethnicity, BMI) $p=0.17$	Information bias		

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
		Students' proxy efficacy in getting peers to create supportive environments for physical activity, two items		
		Students' proxy efficacy in getting schools to create supportive environments for physical activity, six items		
		Students' self-efficacy to be physically active 1–7 days a week, from 0 (not at all sure) to 5 (completely sure), three items		
		Group norms regarding physical activity		
		Students' proxy efficacy in getting parents to create supportive environments for fruit and vegetable consumption, three items		
		Students' proxy efficacy in getting school to create supportive environments for fruit and vegetable consumption, nine items		
		Students' self-efficacy to eat five to seven servings of fruit and vegetables per day, three items		
		Group norms concerning fruit and vegetable consumption		
PeaceBuilders	Flannery <i>et al.</i> 2003 ¹¹⁹	Log-transformed teacher-reported aggressive behaviour among grade 3–5 students, 25 items with three-item response scale, alpha=0.95 at baseline	Aggressive behaviour subscale of the Achenbach Teacher Report Form ¹⁹⁰	Follow-up 1 =spring 1995 (baseline autumn 1994) ¹⁹⁰

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
multilevel mediating variable analysis, which was performed using a mixed model and an approximate standard error test of the mediation effect. No effect estimates were tabulated, only least-squares means and standard errors (SEs) for each measure at each time point by arm with a <i>p</i> -value for the interaction between time and arm. Text reports direction of effect but not size	Not reported	Crude <i>p</i> =0.44; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.48	Information bias	
	Not reported; differences suggested programme benefit	Crude <i>p</i> =0.001; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.001	Information bias	
	Not reported; differences suggested comparison benefit	Crude <i>p</i> =0.02; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.02	Information bias	
	Not reported	Crude <i>p</i> =0.49; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.66	Information bias	
	Not reported	Crude <i>p</i> =0.27; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.28	Information bias	
	Not reported	Crude <i>p</i> =0.71; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.65	Information bias	
	Not reported; differences suggested comparison benefit	Crude <i>p</i> =0.28; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.04	Information bias	
Not reported; differences suggested intervention benefit	Crude <i>p</i> =0.05; adjusted (sex, SES, ethnicity, BMI) <i>p</i> =0.03	Information bias		
Coefficient from hierarchical linear model with three levels – semester, school and individual – for cross-level semester by school interaction effect, adjusting for baseline aggression and sex	Coefficient=0.017	SE=0.005, <i>p</i> <0.01	Information bias as teachers not blinded to arm	Only <i>n</i> =4 schools per arm so potential for random error
Linear regression to examine whether or not regression slopes are non-parallel, indicating that effects are differential by baseline value of the outcome	Not applicable	$t_{1174} = 3.84, p < 0.001$		Moderate rates of dropout although do not appear to be differential by arm

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
		Log-transformed teacher-reported aggressive behaviour among grade K–2 students, 25 items with three-item response scale, alpha=0.95 at baseline		
		Log-transformed child self-reported aggressive behaviour among grade 3–5 students, nine items with three-item response scale, alpha=0.86 at baseline	New measure	
		Log-transformed child self-reported aggressive behaviour among grade K–2 students, nine items with three-item response scale, alpha=0.66 at baseline		
	Krug <i>et al.</i> 1997 ¹²¹	Visits to school nurse for all reasons (weekly rate of visits per 1000 student days)	Routine data from every week	Data from 1994–5, i.e. during intervention delivery; pre-intervention baselines from 1993–4
		Visits to school nurse injuries only (weekly rate of visits per 1000 student days)		
		Visits to school nurse injuries from confirmed fighting (weekly rate of visits per 1000 student days)	Routine data sampled from every fourth week	
		Visits to school nurse non-fighting-related injuries (weekly rate per 1000 student days)		
		Visits to school nurse injuries related to possible fighting (weekly rate per 1000 student days)		
AAYP – SCI	Flay <i>et al.</i> 2004 ⁴⁴	Self-reported violence, multi-item scale used with previous populations modified and piloted for local use	Self-completion surveys in classrooms; trained project staff read questions aloud	Final follow-up spring grade 8 (baseline autumn grade 5)

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
Coefficient (from HLM) as above	Coefficient=0.006	SE=0.005, not significant		Comparison school dropped out at baseline and was replaced with another non-randomised school; therefore, risk of selection bias
Coefficient (from HLM) as above	Coefficient=-0.003	SE=0.01, not significant	Information bias	No data from this replacement comparison school at baseline so potential for residual confounding
Linear regression as above	Not applicable	$t_{1494} = 14.19, p < 0.001$	Information bias	
Coefficient (from HLM) as above	Coefficient=0.02	SE=0.02, not significant	Information bias	
Calculated the average for each school year by giving equal weight to each week regardless of the number of days that the schools was open, to enable comparability between schools	Not reported; differences suggested intervention benefit	$F_{1,247} = 16.43, p < 0.001$	Information bias through intervention affecting recording practices. Interviews with nurses examined alternative explanations and found none	Only four schools per arm to begin with and two schools in comparison excluded because no routine data available
For each outcome calculated the % difference between 1993-4 and 1994-5 by school and wave. Used <i>t</i> -test to determine significance of these differences. Analysis of covariance used to compare control and intervention schools, taking into account baseline differences	Not reported; differences suggested intervention benefit	$F_{1,247} = 42.38, p < 0.001$	Information bias	
	Not reported; differences suggested intervention benefit	$F_{1,55} = 9.06, p = 0.004$	Information bias	
	Not reported; differences suggested intervention benefit	$F_{1,55} = 19.05, p < 0.001$	Information bias	
	Not reported	$F_{1,55} = 0.02, p = 0.89$	Information bias	
Hierarchical statistical models for nested observations (times, subjects, schools). Mixed models – two levels reported as school effects proved negligible. All models included condition, sex, time and all interactions, except for condom use, which was estimated separately for boys (because of low rates of sexual intercourse for girls). Effect size assessed for growth in continuous	Relative reduction = 19%, effect size = 0.10 for boys; relative reduction = 31%, effect size = 0.3 for girls (relative reductions and effect sizes for all outcomes for girls are from personal communication with Brian Flay, 24 October 2011)	$p = 0.52$ for boys, $p = 0.049$ for girls	Information bias as (students not blind to allocation)	Random error as only four schools per arm

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
CAPSLE	Fonagy <i>et al.</i> 2009 ¹²⁰	Self-reported substance use, multi-item scale used with previous populations modified and piloted for local use, ordinal score based on multi-items		
		Self-reported recent sexual intercourse, single ordinal item used with previous populations modified and piloted for local use		
		Self-reported condom use, single item used with previous populations modified and piloted for local use		
		Peer-report aggression, overt and relational, six items; Cronbach's average for all peer-report measures in this study=0.80	Trained researchers administered questionnaires to classrooms in three 15- to 45-minute sessions	T1–T4 (October/November year 1 to March/April year 2)
				T4–T5 (March/April year 2 to October/November year 3)
		Self-report perpetrator of aggression, 5-point Likert scale; Cronbach's average for all self-report measures in this study=0.65		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
scale. Reduction in growth is relative to comparison group. <i>p</i> -values from two-tailed tests. Only <i>p</i> -values reported for girls as there are no programme effects. Effect size is the difference in growth between groups divided by the pooled SD of growth				
As above but generalised estimating equations assessed growth in log odds scale	Relative reduction=4%, effect size=0.03 for boys; relative reduction=13%, effect size=0.2 for girls	<i>p</i> =0.89 for boys, <i>p</i> =0.37 for girls	Information bias	
	Relative reduction=37%, effect size=0.21 for boys; relative reduction=44%, effect size=0.34 for girls	<i>p</i> =0.38 for boys, <i>p</i> =0.28 for girls	Information bias	
	Relative reduction=35%, effect size=0.28 for boys; relative reduction=-57%, effect size=-1.07 for girls	<i>p</i> =0.42 for boys, <i>p</i> =0.03 for girls	Information bias	
HLM analyses longitudinally from T1 across three time points of active intervention, calculating best-fitting slope, adjusted for sex and family income. School-level measures were not associated with outcomes and so were not adjusted for. Scores were converted to T-scores; intervention effects were examined using time by intervention interaction terms. Effect size=beta estimate for main effect/square root of variance estimate at every time point and averaging these to obtain overall effect size for the group	Parameter estimate (effect size)=-0.89 (0.25)	<i>p</i> <0.05	Information bias as (students not blind to allocation)	Only three schools per arm
HLM analyses longitudinally across less intense final year after active intervention ceases	Parameter estimate (effect size)=-0.24 (0.02)	Not significant	Information bias	Did not account for clustering – reported this as unnecessary given small average ICC (=0.04) across outcomes
As above for T1–T4 and T4–T5 respectively	Parameter estimate (effect size)=-0.37 (0.30)	<i>p</i> <0.05	Information bias	
	Parameter estimate (effect size)=-0.29 (0.01)	Not significant	Information bias	

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
		Peer-report victimisation, overt and relational, six items		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
		Self-report victimisation, 5-point Likert scale		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
		Peer-report aggressive bystanding, three items		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
		Peer-report helpful bystanding, three items		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
		Self-report empathy for victims, 5-point Likert scale		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
		Self-report aggression is legitimate attitude, 5-point Likert scale		T1–T4 (October/November year 1 to March/April year 2) T4–T5 (March/April year 2 to October/November year 3)
Playground physical structure intervention	Ridgers <i>et al.</i> 2007 ¹²⁵	Physical activity levels during morning and lunch recess periods	Heart rate telemetry and accelerometry; anthropometry measurements of stature and body mass Recess duration = time from when the school bell rang to start recess to when it rang to conclude recess	Baseline and 6 weeks
	Ridgers <i>et al.</i> 2007 ¹²⁴	Physical activity levels during morning and lunch recess periods	Heart rate telemetry and accelerometry	Assessed across time drawing on baseline data and 6-week and 6-month follow-ups

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
	Parameter estimate (effect size)=-0.97 (0.30)	$p < 0.01$	Information bias	
	Parameter estimate (effect size)=-1.05 (0.07)	Not significant	Information bias	
	Parameter estimate (effect size)=0.22 (0.06)	Not significant	Information bias	
	Parameter estimate (effect size)=0.02 (0.00)	Not significant	Information bias	
	Parameter estimate (effect size)=-0.73 (0.20)	$p < 0.05$	Information bias	
	Parameter estimate (effect size)=0.07 (0.00)	Not significant	Information bias	
	Parameter estimate (effect size)=-2.37 (0.59)	$p < 0.001$	Information bias	
	Parameter estimate (effect size)=-2.53 (0.15)	$p < 0.05$	Information bias	
	Parameter estimate (effect size)= 1.01 (0.26)	$p < 0.01$	Information bias	
	Parameter estimate (effect size)= 1.21 (0.08)	Not significant	Information bias	
	Parameter estimate (effect size)=-0.57 (0.09)	Not significant	Information bias	
	Parameter estimate (effect size)=-0.46 (0.03)	Not significant	Information bias	
Multilevel modelling using three-level measurement: (1) follow-up measurement, (2) pupil and (3) school	MVPA adjusted beta coefficient=4.5; VPA adjusted beta coefficient=1.30	MVPA adjusted: not significant (SE=2.83); VPA adjusted: not significant (SE=0.79)		
Adjustment for baseline recess physical activity, BMI, daily recess time, age, sex	Although there were interactions between the intervention and age and daily recess time ($p < 0.05$) it is not reported what were the effect estimates for these subgroups			
Multilevel modelling using three-level measurement: (1) follow-up measurement, (2) pupil and (3) school	MVPA heart rate adjusted beta coefficient=4.03; VPA heart rate adjusted beta coefficient=2.43	MVPA heart rate: 95% CI 0.15 to 7.91, $p = 0.042$; VPA heart rate: 95% CI 0.06 to 4.80, $p = 0.045$	Students aware of measurement fitted with equipment	
Adjustment for baseline physical activity, BMI, daily recess time, age, sex	MVPA accelerometry adjusted beta coefficient=4.53; VPA accelerometry adjusted beta coefficient=2.32	MVPA accelerometry: 95% CI 0.59 to 8.47, $p = 0.025$; VPA accelerometry: 95% CI 0.71 to 3.93, $p = 0.005$		

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
	Ridgers <i>et al.</i> ¹²⁶	Physical activity levels during morning and lunch recess periods	Heart rate telemetry and accelerometry; established threshold (accelerometry cut-off) used as a measure	Assessed across time drawing on baseline data and 6- and 12-month follow-ups
M-SPAN	Sallis <i>et al.</i> 2003 ¹²²	Physical activity at school, overall (kcal per child per day per school)	Assessors systematically observed physical activity, different times and places (PE, before school, after lunch, after school on school grounds)	Baseline, year 1, year 2
		Physical activity at school, boys (kcal per child per day per school)	System for Observing Fitness Instruction Time used to assess PE; validated measure – 2 randomly selected days per school; interobserver agreement=83%	
		Physical activity at school, girls (kcal per child per day per school)	System for Observing Play and Leisure Activities for Youth; new measure – 2 randomly selected days per school; interobserver agreement=99%	
		MVPA score, boys – intensity values for each physical activity multiplied by minutes to yield weighted score		
		Sedentary hours per day, boys		
		MVPA score, girls, as above		
		Sedentary hours per day, girls		
		Fat intake of average child during school day	Fat on average day summed for each school, adjusted for school size to provide school-level measure.	
		Saturated fat intake of average child during school day	Menu documentation by trained food service staff plus cook interviews.	
		Student-reported fatty foods consumed the previous day, boys	Food labels and sales data collected. For bag lunches trained observers recorded food and serving size. % students bringing bag lunches from student surveys.	
		Student-reported fatty foods consumed the previous day, girls	Students store all food labels collected, sales items recorded	

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
Multilevel modelling using three-level measurement: (1) follow-up measurement, (2) pupil and (3) school	Morning recess: % MVPA (heart rate) risk difference (RD)=4.1; % MVPA (accelerometry) RD=1.2; % VPA (heart rate) RD=3.6; % VPA (accelerometry) RD=1.4	Morning recess: % MVPA (heart rate): 95% CI -1.9 to 10.0; % MVPA (accelerometry): 95% CI -2.8 to 5.2; % VPA (heart rate): 95% CI -0.5 to 7.7; % VPA (accelerometry): 95% CI -0.4 to 3.1	Students aware of measurement (fitted with equipment)	Missing data at 6 and 12 months
Adjustment for sex, age, baseline physical activity, BMI, recess duration	Lunch recess: % MVPA (heart rate) RD=3.0; % MVPA (accelerometry) RD=3.6; % VPA (heart rate) RD=0.9; % VPA (accelerometry) RD=1.4	Lunch recess: % MVPA (heart rate): 95% CI -1.2 to 7.2; % MVPA (accelerometry): 95% CI -0.3 to 7.4; % VPA (heart rate): 95% CI -1.6 to 3.5; % VPA (accelerometry): 95% CI 0.1 to 2.7		
Randomised regression models. For physical activity the covariance structures that maximised Akaike's information criterion (AIC) were compound symmetry for the overall and boys' models and autoregressive for girls. Survey data were aggregated at the school level. Effect sizes (d) were calculated by subtracting the change in control schools from the change in intervention schools and dividing by the pooled SD of change	$d=0.93$	$F_{1,46} = 7.53, p < 0.009$	Information bias as observers not apparently blind to allocation	Strong study
	$d=1.10$	$F_{1,46} = 12.16, p = 0.001$	Information bias	
	$d=0.37$	$F_{1,46} = 0.73, p = 0.396$	Information bias	
	$d=0.09$	$F_{1,46} = 0.04, p = 0.839$	Information bias	
	$d=0.17$	$F_{1,46} = 0.16, p = 0.693$	Information bias	
	$d=0.25$	$F_{1,46} = 0.37, p = 0.548$	Information bias	
	$d=0.11$	$F_{1,46} = 0.14, p = 0.709$	Information bias	
As above but for nutrition the covariance structures that maximised AIC were autoregressive for total fat and compound symmetry for saturated fat	$d=0.03$	$F_{1,46} = 0.01, p = 0.903$	Information bias as intervention may have affected record-keeping by food staff	
	$d=0.13$	$F_{1,46} = 0.08, p = 0.781$		
	$d=-0.13$	$F_{1,46} = 0.10, p = 0.761$		
	$d=-0.03$	$F_{1,46} = 0.006, p = 0.937$		

continued

TABLE 25 Data extraction and quality appraisal of outcome evaluation studies (intervention health outcomes) (continued)

Intervention	Study	Health outcome	Measurement	Follow-up
Playground surface marking	Stratton 2000 ¹¹⁷	BMI, boys	Based on self-reported height and weight	Immediately post intervention
		BMI, girls		
		Parental fat avoidance, boys	Reported by parents using modified version of established scale	
		Parental fat avoidance, girls		
Playground surface marking	Stratton and Mullan 2005 ¹¹⁸	MVPA	Measured by heart rate threshold (radio telemetry)	Immediately post intervention
		VPA		
		Mean heart rate		

SE, standard error.

Form of effect analysis method	Estimate of intervention effect	Estimate of precision	Risk of bias by outcome	Risk of bias overall study
As above	$d=0.83$ $d=-0.12$ $d=-0.03$ $d=0.43$	$F_{1,46}=4.60, p=0.044$ $F_{1,46}=0.09, p=0.771$ $F_{1,46}=0.004, p=0.948$ $F_{1,46}=1.12, p=0.301$	Information bias as students and parents not blind to allocation	
Analysis of variance, analysis of covariance, adjusted for playtime duration	$F_{1,278}=0.18$ crude, $F_{1,272}=0.33$ $F_{1,278}=0.50$ crude, $F_{1,272}=0.48$ $F_{1,278}=0.05$, $F_{1,272}=0.05$	$p>0.05, p>0.05$ $p>0.05, p>0.05$ $p>0.05, p>0.05$		Non-random, $n=1$
Analysis of covariance, adjusted for playtime duration, body mass	$F_{1,204}=13.7$ $F_{1,204}=4.05$	$p<0.01$ $p<0.03$		Non-random, high differential attrition

TABLE 26 Quality appraisal of outcome evaluation studies

Quality criteria	CDP ^{54,55,58,123}	HSE ⁶⁰	HYP ⁶³	PeaceBuilders ^{119,121}
Impact of the intervention reported for all outcomes	Yes	No	Yes	Yes
Random allocation to intervention and comparison conditions	Yes	No	Yes	Yes
Non-random allocation but matching on or adjustment for potential confounders	NA	Yes	NA	NA
Equivalent study groups at baseline	NA ^a	No	NR	NR ^{b,c}
Attrition ≤30% overall	No	Yes	Yes	Yes
Attrition ≤10% difference between groups	No	Yes	Yes	Yes
Analysis accounts for cluster	No	Yes	Yes	Yes
Other issues	Two groups of three schools randomised together to intervention or control. The reporting of participant flow is not clear	Swapping of intended intervention and comparison schools (possible overestimation of benefits)		Comparison school dropped out at baseline and was replaced by another (non-randomised) school with no baseline information. Only 1-year follow-up is relevant as comparisons received intervention thereafter
Blinding of outcome assessment	Yes	NR	NR	NR
Well conducted			✓	
Less well conducted	✓	✓		✓

NA, not applicable; NR, not reported.

a Pre-intervention baselines not measured.

b Intervention effects were tested by examining the time by condition interaction terms, so although baseline differences were not reported they would be adjusted for in the outcome analysis.

c Although not tested for significance, there were substantial differences in confirmed fighting-related injuries between arms reported by Krug *et al.*,¹²¹ but these were accounted for in the outcome analysis.

d Boys receiving the SCI engaged in more violence than boys in the SDC group ($p=0.02$), but this was adjusted for in outcome analysis.

e Analysis based on repeated cross-sectional data not longitudinal data.

AAYP – SCI ⁴⁴	CAPSLE ¹²⁰	Playground physical structure intervention ^{124–126}	M-SPAN ¹²²	Playground surface marking ¹¹⁷	Playground surface marking ¹¹⁸
Yes	Yes	Yes	Yes	Yes	Yes
Yes	Yes	No	Yes	No	No
NA	NA	Yes	NA	Yes (area, playground children's age, size and stature)	Yes (playground dimension, area deprivation)
No ^d	Yes	No	Yes	No – differences in MVPA	Yes
NA ^e	No	Not clear	NA ^e	Yes	No
NA ^e	No	Not clear	NA ^e	No	No
Yes	No One school allocated to comparison dropped out before the intervention and was not replaced	Yes	Yes	No	No
NR	NR	NR	NR	NR	NR
✓	✓	✓	✓	✓	✓

Research question 3: process evaluation

TABLE 27 Data extraction of process evaluation studies

Study	Process evaluation methods	Process evaluation findings
Battistich <i>et al.</i> 1989 ⁵⁵	Delivery – each class visited eight times per school year for 2 hours by observers blind to arm; structured instrument. Data aggregated to yield implementation score for each class for each year. Inter-rater reliability=0.54 overall	Programme classrooms across the 5 years had significantly higher mean implementation scores for each of the programme components [multivariate analysis of variance (MANOVA) $F_{5,61}=13.67$, $p<0.0001$; univariate t 's $65>3.06$, p 's <0.01]
Bonell <i>et al.</i> 2010 ⁵⁹	Baseline ($n=605$) and follow-up surveys ($n=721$) with year 7 students in two intervention/ comparison schools Semistructured interviews with staff ($n=15$), students ($n=45$) and facilitators ($n=2$) Semistructured observations ($n=13$)	The intervention was delivered as intended with all components implemented. The external facilitator enabled schools to convene an action team involving staff/students. Inputs were feasible and acceptable and enabled similar actions in both schools. Locally determined actions (e.g. peer mediators) were generally more feasible and acceptable than preset actions (e.g. modified pastoral care). Implementation was facilitated when it built on aspects of schools' baseline ethos (e.g. a focus on engaging all students, formalised student participation in decisions) and when senior staff led actions. Student awareness of the intervention was high
Bonell <i>et al.</i> 2010 ⁶⁰	Baseline ($n=614$) and follow-up surveys ($n=735$) with year 7 students in two intervention/ comparison schools Semistructured interviews with staff ($n=4$), students ($n=42$) and facilitators ($n=2$)	Student accounts linked participation in planning or delivering intervention activities with improved self-regard and relationships with staff and other students. Some activities, such as rewriting school rules, involved broad participation. Students in receipt of actions such as peer mediation or motivational sessions reported benefits such as improved safety and relationships. Some student accounts linked improved self-regard and relationships with increased engagement and aspirations and reduced substance use
Dzewaltowski <i>et al.</i> 2009 ⁶³	Site co-ordinator training – attendance and survey, assessing self-efficacy to lead and train others to build healthy places Intervention delivery – site co-ordinators self-reporting meetings, logging implemented programmes, policies and environmental changes, evaluated by two independent raters Curriculum – teachers' self-reported implementation	Training – attendance was 97% in the first intervention year and 91% in the second intervention year. Site co-ordinators' self-efficacy to lead and train others to implement the intervention was high (4.00 on a 0–5 scale; SD = 0.45) at the beginning of the project and did not change after year 2 (3.79, SD=1.05) and year 3 (4.10, SD=0.66) Delivery – site co-ordinators formed changed teams during the first intervention year and began meeting regularly during the spring. During year 2 the site co-ordinators held an average of 15.8 meetings. The site co-ordinators reported an average of 26.5 implemented programme, policy or practice changes Curriculum – during seventh grade teachers reported implementing 5.6 lessons in the classroom for 13.5 hours, which was on average 64.4% of the intended lessons, but implemented lessons exceeded the lesson time goal (mean 122.6%). During eighth grade school teachers reported implementing 3.25 lessons in the classroom for 5.89 hours, which was on average 81.2% of the intended lessons and 60.1% of the intended lesson time goal

TABLE 27 Data extraction of process evaluation studies (continued)

Study	Process evaluation methods	Process evaluation findings
	Intervention receipt – student survey of awareness and participation	Receipt – 31.6% of students had heard about HYP. In total, 35.6% perceived that there were media promoting physical activity after school, 15.2% perceived that there were media promoting fruit and vegetable consumption, 31.0% thought that school staff promoted physical activity after school and 14% thought that school staff promoted fruit and vegetable consumption. Just over one-third of students (37.5%) had heard about the change team, 14.9% participated on the change team, 23.4% had heard about the video team and 7.2% participated on the video team
Flannery <i>et al.</i> 2003 ¹¹⁹	Teacher training – 10-item survey assessing clarity and effectiveness and impressions of whether materials and programme would be easy or difficult to implement, rated on a 5-point scale. Response rates from 75% to 86%	Training – all teachers participated ($n=194$). Overall, 93% of teachers indicated that they strongly agreed or agreed that the basic philosophy behind the PeaceBuilders intervention was easy to understand; 77% agreed/strongly agreed that the training provided was clear, effective and easy to follow; and 83% agreed/strongly agreed that the ideas would be easy to use in the classroom. Three of four teachers who completed surveys believed that 'PeaceBuilders will be very successful as an intervention' and strongly agreed or agreed that 'the school administration stands behind this intervention effort 100 percent'
	Implementation and fidelity – in the spring of year 2, teachers completed an eight-item survey assessing their use and implementation of programme materials, including frequency, satisfaction and effectiveness. Teachers were also asked to indicate the number of PeaceBuilders materials used in their classrooms	Implementation and fidelity – 190 teachers (98%) completed a spring 1996 self-assessment of their use of intervention materials in the classroom. A majority of teachers surveyed indicated that they used the PeaceBuilders curriculum in their classrooms on a daily (48%) or weekly (32%) basis. Nearly all teachers (98%) strongly agreed or agreed that, 'overall, my school has implemented the PeaceBuilders curriculum', 53% rated implementation as 'extensive' and 43% rated implementation as 'moderate'. Regarding the total number of programme materials used, teachers reported, on average, that they used at least four of the eight core sets of materials in their classrooms. Teachers in the control schools reported, more than did teachers in the intervention schools, that during year 2 they were more likely to use programme materials daily than weekly ($p<0.01$)
Solomon <i>et al.</i> 1988 ⁵⁸	Observed delivery – in each of the 5 study years classrooms in all six schools were observed during eight separate 2- to 5-hour visits (95 classrooms over the 5 years, approximately half in programme schools and half in comparison schools). Observers were trained and were not aware of allocation	Observations – combining all grades there were significantly more observations of the following in intervention classrooms: co-operative activities ($p<0.001$), developmental discipline ($p<0.01$), social understanding ($p<0.001$), prosocial values ($p<0.001$), helping activities ($p<0.001$) and on an overall implementation index ($p<0.001$)
	Teacher reports of delivery and context – teacher questionnaires each spring, asking for descriptions of classroom events and activities during the year, goals in teaching and discipline, perceptions of school climate and assessments of their students' behaviour	Teacher reports – there were significant differences between intervention and control arms (former higher unless otherwise stated) in time students in groups spent doing independent work with help from other students ($p<0.01$); student participation in rule development ($p<0.001$); problem-solving class

continued

TABLE 27 Data extraction of process evaluation studies (continued)

Study	Process evaluation methods	Process evaluation findings
	<p>Student reports of delivery – in the spring of the third-grade year, students were asked, in individual interviews, questions about their perceptions of classroom activities, the reasons for doing them and teachers' goals and intentions</p>	<p>discussions ($p < 0.001$); students working out own methods of learning and problem-solving ($p < 0.01$); social understanding discussions ($p < 0.01$); teacher use of individual rewards ($p < 0.01$), with control arm higher; teacher use of group praise ($p < 0.1$); teacher use of group rewards ($p < 0.1$, control arm higher). For nine other measures there were no significant differences ($p < 0.1$)</p> <p>Student reports – there were significant differences between intervention and control arms (former higher unless otherwise stated) in teacher likes students to work well with others ($p < 0.05$); teachers likes students to be helpful, nice to others ($p < 0.05$); students made the class rules ($p < 0.01$); both teachers and students made the class rules ($p < 0.01$); punishment if student breaks rule ($p < 0.01$, control higher); warning, name on board if student breaks rule ($p < 0.05$, control higher); temporary separation, time out if student breaks rule ($p < 0.01$); reparation if student breaks rule ($p < 0.01$); discuss with teacher/other adults if student breaks rule ($p < 0.1$); rule would change by student getting teacher/principal to change it ($p < 0.05$); do academic work in groups ($p < 0.01$); teacher uses group work so work will be carried out better ($p < 0.05$, control higher); teacher uses group work so learn to be co-operative ($p < 0.01$); teacher uses group work so learn to understand/appreciate others ($p < 0.01$); and teacher uses group work to prepare for adulthood ($p < 0.01$). For eight other measures there were no significant differences ($p < 0.1$)</p>

TABLE 28 Quality appraisal of process evaluation studies

Study	Clear RQ	Sampling described	Study sample	Examined planning qualitative?	Examined delivery?	
					Quantitative	Qualitative
Battistich <i>et al.</i> 1989 ⁵⁵	No	Yes (classes)	NA	No	Observed fidelity of teaching practices	No
Bonell <i>et al.</i> 2010 ⁵⁹	Yes	Yes (interviews but not observations)	Yes	Yes	Recorded fidelity – actions taken	Yes – feasibility
Bonell <i>et al.</i> 2010 ⁶⁰	Yes	Yes (interviews)	Yes	Yes	No	No
Dzewaltowski <i>et al.</i> 2009 ⁶³	No	Yes (students); no (site co-ordinators, teachers)	Yes (students); no (site co-ordinators, teachers)	No	Fidelity – site co-ordinator reported training delivery, number of meetings and number of policy changes; teacher reported 'curriculum' delivery	No
Flannery <i>et al.</i> 2003 ¹¹⁹	No	Yes (teachers)	Yes (teachers)	No	Fidelity – teacher reported delivery	No
Solomon <i>et al.</i> 1988 ⁵⁸	Yes	Yes (observations, teachers, students)	No	No	Fidelity – observed, teacher- and student-reported classroom teaching practices	No

NA, not applicable.

Examined coverage?		Examined receipt?		Examined acceptability?		Examined context?	
Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative
No	No	No	No	No	No	No	No
Student awareness	Yes	No	Yes	No	Yes (students, teachers)	Yes, local need, school capacity, staff capacity, product champion, collaboration	Yes
Student awareness	No	No	Yes	No	Yes (students, teachers)	Yes, measured need	Yes
Student recognition	No	Site co-ordinator self-efficacy to lead	No	No	No	No	No
No	No	Teacher reported acceptability of training, programme	Yes (teachers)	Yes	No	No	No
No	No	No	No	No	No	No	No

Research question 4: multilevel studies

TABLE 29 Data extraction of multilevel studies (study context)

Study	Year of data collection	Data set (if mentioned)	Location/country	Type of school(s)
Adams <i>et al.</i> 2009 ¹³⁴	Not stated	Youth Tobacco Access Project	USA, (IL)	Middle and high schools
Akiba 2010 ⁷⁰	2000	Program for International Student Assessment	USA	Not stated
Attar-Schwartz and Khoury-Kassabri 2008 ¹³⁸	2005		Israel	Jewish and Arab secondary schools
Aveyard <i>et al.</i> 2004 ⁴⁸	1995/6	West Midlands Young People's Lifestyle Survey	UK (West Midlands)	Secondary and middle schools
Bisset <i>et al.</i> 2007 ⁶⁶	1995–6	West Midlands Young People's Lifestyle Survey	UK (West Midlands)	Secondary schools
Bradshaw <i>et al.</i> 2009 ⁸⁹	2005–6		USA (MD)	Elementary and middle public schools
Collins <i>et al.</i> 2008 ⁷²	1999–2002		USA (KY)	Schools with eighth graders
Comeau 2005 ⁸⁵	Not stated	Next Generation Project	USA (Eugene, OR)	Middle schools
Cradock <i>et al.</i> 2007 ¹³³	1997		USA (Boston metropolitan area, MA)	Middle schools
Evans-Whipp <i>et al.</i> 2010 ¹³¹	2003	International Youth Development Study	USA (WA), Australia (VIC)	Secondary schools
Gastic 2005 ⁷⁶	1994 and 1996	NLSAH 'Add Health'	USA (west, midwest, south and north-east regions)	High schools (grades 9–12) and middle schools (grades 7–8)
Gladden 2005 ⁸⁶	1997 and 1999	Consortium on Chicago Schools Research	USA (Chicago, IL)	Public elementary schools
Grunseit <i>et al.</i> 2008 ⁹⁷	Not stated		Australia (NSW)	Secondary schools
Haley 2007 ⁷⁷	1994–5	NLSAH ('Add Health')	USA	High schools
Haug <i>et al.</i> 2010 ⁸⁷	2005	Health Behaviour in School-aged Children	Norway	Secondary schools
Henry <i>et al.</i> 2009 ⁶⁷	1996–2000		USA	Public high schools, junior high/middle schools
Ishibashi 2009 ⁸³	2005–6	SHAPES	Canada (ON)	Secondary schools (grades 9–12)
Johnson and Hoffman 2000 ⁷¹	1988 and 1990	NELS	USA	Catholic, other private, public schools
Kairouz and Adlaf 2003 ⁷⁹	1999	OSDUS	Canada (ON)	High schools

TABLE 29 Data extraction of multilevel studies (study context) (continued)

Study	Year of data collection	Data set (if mentioned)	Location/country	Type of school(s)
Khoury-Kassabri <i>et al.</i> 2009 ⁷⁸	2005		Israel	Junior high and high public schools
Kim 1999 ⁸¹	1988 and 1990	NELS	USA (nationally representative)	Public and private high schools
Kim 2001 ⁸⁸	Not stated	NLSAH ('Add Health')	USA	High schools and schools that feed into high schools
Kumar <i>et al.</i> 2008 ⁶⁸	2000–3	Monitoring the Future Project	USA	Public and private schools
Leatherdale <i>et al.</i> 2010 ⁹⁰	2007–8	'Play Ontario' trial	Canada (ON)	Elementary schools (grades 5–8)
Lothian 2003 ⁸⁰	1999	OSDUS	Canada (ON)	Elementary and secondary schools
Ma 2000 ¹³⁵	1996	New Brunswick School Climate Study	Canada (NB, Atlantic Province)	English elementary schools
Maes and Lievens 1999 ¹⁴⁰	1994	Extension of Flemish Health Behaviour School-aged Children Survey	Belgium	Secondary
Malikaew <i>et al.</i> 2003 ¹⁴²	Not stated		Thailand (Muang District, Chiang Mai Province)	Primary schools
Markham <i>et al.</i> 2008 ⁶⁵	1997 and 1999		UK (West Midlands)	Secondary schools
Monshouwer <i>et al.</i> 2007 ⁶⁹	2003	Dutch National School Survey	Netherlands	Secondary schools (pre-vocational, lower, upper, pre-university)
Moore <i>et al.</i> 2001 ¹³⁶	1998	Welsh Youth Health Survey (part of International Health Behaviour School Age Children)	Wales	Secondary
Murnaghan <i>et al.</i> 2007 ⁸⁴	1999–2001	SHAPES	Canada (Prince Edward Island)	English-speaking secondary schools (grades 10–12)
Pinilla <i>et al.</i> 2002 ¹⁴¹	2000		Spain (Gran Canaria)	Secondary
Piontek <i>et al.</i> 2008 ¹³²	Not stated		Germany (Bavaria)	Secondary schools – Hauptschulen (low level of education), Realschulen (intermediate level), Gymnasien (high level), Gesamtschulen (intermediate and high level)
Rountree and Clayton 1999 ⁷³	1996	Kentucky Youth Survey	USA (KY)	Public schools with grades 6–12
Tobler <i>et al.</i> 2011 ⁵³	2002, 2003, 2004 and 2005		USA (Chicago, IL)	Public schools with grades 5–8
van der Horst <i>et al.</i> 2009 ¹³⁹	2005–6	Environmental Determinants of Obesity in Rotterdam School children	Netherlands (Rotterdam)	Vocational schools and higher-level secondary education schools

continued

TABLE 29 Data extraction of multilevel studies (study context) (continued)

Study	Year of data collection	Data set (if mentioned)	Location/ country	Type of school(s)
Way 2003 ⁸²	1988, 1990, 1992 and 1994	NELS	USA (nationally representative)	Private, public and Catholic high schools
West <i>et al.</i> 2004 ¹³⁰	1994, 1996 and 1999	West of Scotland 11–16 Study	UK (Central Clydeside conurbation, Scotland)	Primary and secondary schools
West <i>et al.</i> 2010 ¹³⁷	1994, 1996, 1999 and 2002–3	West of Scotland 11–16 Study	UK (Scotland)	Primary and secondary schools (public and private)
Wium and Wold 2006 ⁷⁴	Not stated	CAS study	Norway	Schools with tenth graders
Wium and Wold 2009 ⁷⁵	Not stated	CAS study	Norway	Schools with tenth graders

TABLE 30 Data extraction of multilevel studies (response and follow-up rate)

Study	School response rate at baseline	Student response rate at baseline	Waves of follow-up post baseline	School response rate at each follow-up	Student response rate at each follow-up
Adams <i>et al.</i> 2009 ¹³⁴	40 (denominator not reported)	16,676/30,969 (54%), but 115 (0.7%) excluded because of inconsistent or invalid responding	NA		
Akiba 2010 ⁷⁰	111 (denominator not reported)	2787 (denominator not reported)	NA		
Attar-Schwartz and Khoury-Kassabri 2008 ¹³⁸	324 (95%)	16,604 (denominator not reported)	NA		
Aveyard <i>et al.</i> 2004 ⁴⁸	166 (62%)	23,282 (denominator not reported)	NA		
Bisset <i>et al.</i> 2007 ⁶⁶	201/329 (61%), 166 secondary (not mid) schools in this sample	27,257 (91%), 25,789 students from this sample	NA		
Bradshaw <i>et al.</i> 2009 ⁸⁹	95/95 (100%)	22,178 (74%)	NA		
Collins <i>et al.</i> 2008 ⁷²	59 (denominator not reported)	22,986 (denominator not reported)	NA		
Comeau 2005 ⁸⁵	8 (denominator not reported)	1076/1106 giving consent (97%)	NA		
Cradock <i>et al.</i> 2007 ¹³³	10 (denominator not reported)	248 (denominator not reported)	NA		
Evans-Whipp <i>et al.</i> 2010 ¹³¹	285 (no denominator)	3466 (no denominator)	NA		
Gastic 2005 ⁷⁶	132 (100%)	Not stated	NA		
Gladden 2005 ⁸⁶	323 (68%) in 1997 and 283 57% in 1999	17,700 (78%)	One in this analysis		'Over 15,000' (79%)
Grunseit <i>et al.</i> 2008 ⁸⁷	60/69	2616 (80.5%)	NA		
Haley 2007 ⁷⁷	132 (79%)	18,047 (denominator not reported)	NA		
Haug <i>et al.</i> 2010 ⁸⁷	68/115	1347/2754	NA		
Henry <i>et al.</i> 2009 ⁶⁷	349 (denominator not reported)	43,465 (response rates said to vary from 75% to 100%)	NA		

continued

TABLE 30 Data extraction of multilevel studies (response and follow-up rate) (continued)

Study	School response rate at baseline	Student response rate at baseline	Waves of follow-up post baseline	School response rate at each follow-up	Student response rate at each follow-up
Ishibashi 2009 ⁸³	76 (64%) schools from 19 (86%) school boards	51,739 students (81.7%)	NA		
Johnson and Hoffmann 2000 ⁷¹	1012 (denominator not reported)	16,454 (92%)	One	1397 (denominator not reported)	13,840 (92%)
Kairouz and Adlaf 2003 ⁷⁹	Stated 72/72	3330 (denominator not reported)	NA		
Khouny-Kassabri <i>et al.</i> 2009 ⁷⁸	324 (denominator not reported)	16,604 (denominator not reported)	NA		
Kim 1999 ⁸¹	1035 (denominator not reported)	18,270 (denominator not reported)	One	1035 (denominator not reported)	18,270 + 1043 (denominator not reported)
Kim 2001 ⁸⁸	123 (denominator not reported)	12,604 (denominator not reported)	NA		
Kumar <i>et al.</i> 2008 ⁶⁸	Not stated	27,462 eighth graders + 21,920 tenth graders + 21,501 twelfth graders (denominator not reported)	NA		
Leatherdale <i>et al.</i> 2010 ⁹⁰	30/30 (100%)	2449/4838 (50.6%)	NA		
Lothian 2003 ⁸⁰	111/123	12,814 (76%)	NA		
Ma 2000 ¹³⁵	147 (100%)	6883 ('nearly 100%')	NA		
Maes and Lievens 1999 ¹⁴⁰	NR	NR			
Malikaew <i>et al.</i> 2003 ¹⁴²	52/53	2725 (denominator not reported)	NA		
Markham <i>et al.</i> 2008 ⁶⁵	52/89	8352 (92.0%)		52 and 50	7444 (89.1% of baseline); 6819 (84.6% of baseline)
Moore <i>et al.</i> 2001 ¹³⁶	55/80	NR			
Monshouwer <i>et al.</i> 2007 ⁶⁹	192 (72%)	7324 (93%)	NA		
Murnaghan <i>et al.</i> 2007 ⁸⁴	10/10	3965 (denominator not reported)	NA		

Study	School response rate at baseline	Student response rate at baseline	Waves of follow-up post baseline	School response rate at each follow-up	Student response rate at each follow-up
Pinilla <i>et al.</i> 2002 ¹⁴¹	30/33	NR			
Piontek <i>et al.</i> 2008 ¹³²	40/48	3364 (no denominator)	NA		
Rountree and Clayton 1999 ⁷³	53/53	2295/3359 (68%)	NA		
Tobler <i>et al.</i> 2011 ⁵³	61 (denominator not reported)	<i>n</i> not reported (response rates ranged from 91% to 96% per year)	Three	School response rate at each follow-up not stated	2621 at follow-up 2 (61%)
van der Horst <i>et al.</i> 2009 ¹³⁹	16/17	1206/1668	NA		
Way 2003 ⁸²	1052 (denominator not reported)	'Over 25,000' (denominator not reported)	Three	Not reported in full but 877/1052 by final follow-up	Not reported in full but 13,371 by final follow-up
West <i>et al.</i> 2004 ¹³⁰	43/44	2583 (93%)	Two	43/43 and 39/43 (four schools closed and replaced by four new schools to which most of the students transferred)	2371 (85%) and 2196 (79%)
West <i>et al.</i> 2010 ¹³⁷	43/44	2586 (93%)	Three	43/43 and 39/43 (four schools closed and replaced by four new schools to which most of the students transferred)	2371 (85%) and 2196 (79%) and 1258 (45%)
Wium and Wold 2006 ⁷⁴	73/89	1404/2156	NA		
Wium and Wold 2009 ⁷⁵	73 (denominator not reported)	1404 (65.2%)	NA		

NA, not applicable; NR, not reported.

TABLE 31 Data extraction of multilevel studies (design and sampling)

Study	Design	Data (year)	Location	Age (years)	School sampling strategy	No. of schools sampled	Mean no. of students sampled per school
Adams <i>et al.</i> 2009 ¹³⁴	C	NR	IL, USA	12–18	NR	20 middle + 20 high = 40	414
Akiba 2010 ⁷⁰	C	2000	USA	15	Three-stage stratified sample [Organisation for Economic Cooperation and Development Programme for International Student Assessment (OECD PISA)]	111	25
Attar-Schwartz and Khouny-Kassabri 2008 ¹³⁸	C	2005	Israel	12–17	Two-stage stratified random cluster sample	324	51
Aveyard <i>et al.</i> 2004 ⁴⁸	C	1995–6	West Midlands, UK	11–12, 13–14, 15–16	Stratified random cluster sample	166	140
Bisset <i>et al.</i> 2007 ⁸⁶	C	1995–6	West Midlands, UK	11–12, 13–14, 15–16		166	
Bradshaw <i>et al.</i> 2009 ⁸⁹	C	2005–6	MD, USA	9–14	Attempted census of a large public school district	95 (76 elementary, 19 middle)	233
Collins <i>et al.</i> 2008 ⁷²	C	2002	KY, USA	13–14	NR	159	145
Comeau 2005 ⁸⁵	Not clear	NR	School District 4J, Eugene, OR, USA	11–14	Attempted census	8	134
Cradock <i>et al.</i> 2007 ¹³³	C	1997	Boston metropolitan area, MA, USA	12–14	Stratified random sample	10	25

Study	Design	Data (year)	Location	Age (years)	School sampling strategy	No. of schools sampled	Mean no. of students sampled per school
Evans-Whipp <i>et al.</i> 2010 ³¹	C	2003	VIC, Australia, and WA, USA	11–12, 13–14, 15–16	State-wide representative samples	285	12
Gastic 2005 ⁷⁶	C	1994–6 (for waves 1 and 2, but only wave 1 used for health outcome model)	USA	12–18	Nationally representative sample of schools	100	43
Gladden 2005 ⁸⁶	C	1997 and 1999	Chicago, IL, USA	13–14	Attempted census of Chicago Public Schools	283	53
Gruseit <i>et al.</i> 2008 ⁹⁷	C	NR	NSW, Australia	13–15	Two-stage stratified cluster sample	60	36
Haley 2007 ⁷⁷	C	1994–5	USA	12–18 same sex-attracted subgroup	Stratified random sample	125	145
Haug <i>et al.</i> 2010 ⁸⁷	C	2005	Norway	13	Cluster sample of one class per school	68	20
Henry <i>et al.</i> 2009 ⁶⁷	C	1996–2000	USA	13–14	Stratified representative sample of rural schools in the contiguous USA, oversample of ethnic minorities	349	125
Johnson and Hoffman 2000 ⁷¹	L	Eighth-grade panel: wave 1: 1988, wave 2: 1990; tenth-grade panel: wave 1: 1990, wave 2: 1992	USA	Eighth-grade panel: wave 1: 13–14, wave 2: 15–16; tenth-grade panel: wave 1: 15–16, wave 2: 17–18	Two-stage probability sample	Eighth-grade panel: 1012; tenth-grade panel: 1397	Eighth-grade panel: 16; tenth-grade panel: 10
Kairouz and Adlaf 2003 ⁷⁹	C	1999	ON, Canada	14–18 (outliers at 13 and up to 20)	Two-stage stratified cluster sample	72	41
Khoury-Kassabri <i>et al.</i> 2009 ⁷⁸	C	2005	Israel	12–17	Non-proportional two-stage stratified cluster sample	324	51
Kim 1999 ⁸¹	C	1988 and 1990 first and second wave data	USA	Data from ages 13–14 and 15–16	Two-stage stratified cluster sample	Unclear	Unclear
Kim 2001 ⁸⁸	C	NR	USA	12–18	Systematic stratified sample	123	102

continued

TABLE 31 Data extraction of multilevel studies (design and sampling) (continued)

Study	Design	Data (year)	Location	Age (years)	School sampling strategy	No. of schools sampled	Mean no. of students sampled per school
Kumar <i>et al.</i> 2008 ⁶⁸	C	2000–3	MI, USA	13–14, 15–16, 17–18	Nationally representative sample	Eighth grade: 244; tenth grade: 211; twelfth grade: 200	Eighth grade: 113; tenth grade: 104; twelfth grade: 108
Leatherdale <i>et al.</i> 2010 ⁹⁰	C	2007–8	ON, Canada	10–14	Convenience sample	30	79
Lothian 2003 ⁸⁰	C	1999	ON, Canada	12–18	Two-stage stratified sample plus additional data from three school boards in Ontario	198	65
Ma 2000 ¹³⁵	C	1996	NB, Canada	11–12	Census	147	47
Maes and Lievens 1999 ¹⁴⁰	C	1994	Belgium	12–13 and 16–17	Purposive	29	111
Malikaew <i>et al.</i> 2003 ¹⁴²	C	NR	Muang District, Chiang Mai Province, Thailand	12	Census	52	52
Markham <i>et al.</i> 2008 ⁶⁵	L	1997–9	West Midlands, UK	13–14 followed up 1 and 2 years later	Random sample probability proportional to size	1997: 52; 1998: 52; 1999: 50	1997: 161; 1998: 143; 1999: 136
Moore <i>et al.</i> 2001 ¹³⁶	C	1998	Wales, UK	15–16	Random	55	25
Monshouwer <i>et al.</i> 2007 ⁶⁹	C	2003	Netherlands	12–16	Two-stage stratified cluster sample	192	38
	C	1999–2001		17–18	NR	10	397

Study	Design	Data (year)	Location	Age (years)	School sampling strategy	No. of schools sampled	Mean no. of students sampled per school
Murnaghan <i>et al.</i> 2007 ⁸⁴			Prince Edward Island, Canada				
Pinilla <i>et al.</i> 2002 ¹⁴¹	C	2000	Spain	13–14	Random	30	63
Piontek <i>et al.</i> 2008 ¹³²	C	NR	Bavaria, Germany	10–21	Intervention evaluation pilot and matched schools selected, random sample of classes drawn	40	70
Rountree and Clayton 1999 ⁷³	C	1996	KY, USA	11–18	Particular counties purposively sampled and all schools with relevant ages included	53	43
Tobler <i>et al.</i> 2011 ⁵³	L	2002–5	Chicago, IL, USA	Sixth grade at baseline followed up to grade 8	Purposive for RCT – include grades 5–8, low mobility, 30+ students per grade	61	42.97 in cohort
van der Horst <i>et al.</i> 2009 ¹³⁹	C	2005–6	Rotterdam, Netherlands	12–15	Two-stage stratified sample (report a bit confusing on school selection)	16	75
Way 2003 ⁸²	L	1988, 1990, 1992, 1994	USA	13–20	Two-stage probability sample	877	15
West <i>et al.</i> 2004 ¹³⁰	L	Wave 1: 1994–5; wave 2: 1995–6; wave 3: 1998–9	Glasgow, Scotland, UK	Wave 1: 11; Wave 2: 13; Wave 3: 15	Stratified sample of secondary schools; all primary schools feeding them then sampled, and classes and pupils within them	43 secondary schools, 135 primary schools	Wave 1: 19; wave 2: 55; wave 3: 51; wave 4: 29

continued

TABLE 31 Data extraction of multilevel studies (design and sampling) (continued)

Study	Design	Data (year)	Location	Age (years)	School sampling strategy	No. of schools sampled	Mean no. of students sampled per school
West <i>et al.</i> 2010 ¹³⁷	L	Wave 1: 1994–5; wave 2: 1995–6; wave 3: 1998–9; wave 4: 2002–3	Glasgow, Scotland, UK	Wave 1: 11; wave 2: 13; wave 3: 15; wave 4: 18–19	Stratified sample of secondary schools; all primary schools feeding them then sampled, and classes and pupils within them	43 secondary schools, 135 primary schools	Wave 1: 19; wave 2: 55; wave 3: 51; wave 4: 29
Wium and Wold 2006 ⁷⁴	C	1998	Norway	15	Stratified systematic sample of schools	73	19
Wium and Wold 2009 ⁷⁵	C	1998	Norway	15	Stratified systematic sample of schools	73	19

C, cross-sectional; L, longitudinal; NR, not reported.

TABLE 32 Data extraction of multilevel studies (health outcomes)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Adams <i>et al.</i> 2009 ¹³⁴	Observation of minors using tobacco on school grounds on at least 1 day in the last week (attitudinal, self-report, dichotomy, not confidential)	NR	NR	Sex, school grade, race	Smoking in past 30 days (behavioural, self-report, dichotomy, not confidential)
	Smoking in past 30 days (behavioural, self-report, dichotomy, not confidential)	NR	NR	Sex, school grade, race	Smoking in past 30 days (behavioural, self-report, dichotomy, not confidential)
Akiba 2010 ⁷⁰	Fear of school violence index based on five items, (attitudinal, self-report, interval)	1.49 (0.55) [not % prevalence: interval index range from 1 (low) to 4 (high)] % prevalence of agree/strongly agree answers to index items: 9.9, 7.1, 4.9, 3.9, 4.1	0.047	Sex, race, school mean parental education, rural/urban location (compared with reference of suburban)	Classroom disorder, clear and fair rules, academic achievement, student belonging, student-teacher bonding, perceived teacher support – all measured at the respondent level by pupil self-report
Attar-Schwartz and Khoury-Kassabri 2008 ¹³⁸	Verbal victimisation (index based on four dichotomous items experienced at least once in the previous month, experiential, self-report, confidential)	2.00 (1.47) [not % prevalence: interval index range 0–4 (items summed)], 77.3% reported at least one item	0.2103	Sex, grade, school-level % families with low income, school-level % families with low education, school ethnic affiliation	School policy, teacher support, students' participation, school size, class size – because these are all measured at the respondent level by pupil self-report
	Indirect victimisation index based on four dichotomous items experienced at least once in the previous month, experiential, self-report, confidential)	1.11 (1.14) [not % prevalence: interval index range 0–4 (items summed)], 62.1% reported at least one item	0.0971	Sex, grade, school-level % families with low income, school-level % families with low education, school ethnic affiliation	School policy, teacher support, students' participation, school size, class size – because these are all measured at the respondent level by pupil self-report

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Enforcement of school tobacco policy	OR=0.49	$p < 0.01$	NR	Large loss of sample due to parental permission. Final sample was 53.5% of eligible 30,969 = 16,561. Self-report of smoking carried out in small groups of peers
Enforcement of school tobacco policy	OR=0.83	$p < 0.05$	NR	
Academic tracking	Unstandardised coefficient = 0.049 (SE = 0.024); standardised coefficient = 0.038; Friedman's effect size = 0.076	$p < 0.05$	0.85 (but school-level variance was only 5% of total)	Item response theory was used to impute values for academic achievement as respondents only take a subset of test items in the Programme for International Student Assessment data. A sampling weight was used. Parental education was included both at level 1 and as a mean for the school at level 2 The reported SE (0.001) must be wrong for level of instructional support as it would be reported as highly significant if so. It must actually be 0.010? Either the SE is wrong or the reporting of significance level is wrong
Level of instructional support	Unstandardised coefficient = 0.010 (SE = 0.001); standardised coefficient = 0.026; Friedman's effect size = 0.052	$p > 0.05$		
Fully certified teachers	Unstandardised coefficient = -0.003 (SE = 0.013); standardised coefficient = -0.003; Friedman's effect size = -0.006	$p > 0.05$		
School size	Coefficient = 0.054	$p > 0.05$	0.5679	Sample weights were used to make sample representative of the Israeli student population
Class size	Coefficient = -0.012	$p > 0.05$		
School size	Coefficient = 0.019	$p > 0.05$	0.4011	
Class size	Coefficient = 0.011	$p > 0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Aveyard <i>et al.</i> 2004 ⁴⁸	Regular smoking – whether regularly smoking at least one cigarette per week (data were checked across multiple items for consistency – see reviewer’s comments; dichotomous, behavioural, self-report, anonymous)			School year, sex, age within school year, ethnicity, SES (housing tenure, free school meal status, Townsend index), smoking habits of mother, father and siblings	
	Regular smoking – whether regularly smoking at least one cigarette per week (data were checked across multiple items for consistency – see reviewer’s comments; dichotomous, behavioural, self-report, anonymous)			School year, sex, age within school year, ethnicity, SES (housing tenure, free school meal status, Townsend index), smoking habits of mother, father and siblings	
Bisset <i>et al.</i> 2007 ⁶⁶	Drinking alcohol at least once a month (analysis limited to grade 7 pupils) (behavioural, dichotomous, self-report)			Sex, ethnicity, SES (housing tenure, free school meal status, Townsend index), whether or not drinks with parents	
	Drinking alcohol at least once a month (analysis limited to grade 7 pupils) (behavioural, dichotomous, self-report)				
	Heavy drinking – drinking ≥ 10 units of alcohol per week (behavioural, dichotomous, self-report)			Grade, sex, ethnicity, SES (housing tenure, free school meal status, Townsend index), whether or not drinks with parents	
	Heavy drinking – drinking ≥ 10 units of alcohol per week (behavioural, dichotomous, self-report)				

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School achievement: five grade A–C GCSEs, based on 5-year average (1994–8) % of pupils achieving those grades	OR=1.01 (for 10% increase in rate of five grade A–C GCSEs)	95% CI 0.95 to 1.07 (for 10% increase in rate of five grade A–C GCSEs)	NR	Only 62% school response rate; additionally, 31 schools left out because of not taking GCSEs; inconsistent answers to smoking question items also resulted in removal from data set. Models contained random intercept for school identifier and random effects for school years
School truancy: based on 5-year average (1994–8) of % of half-days lost through pupil absence	OR=1.01 (for 1% increase in truancy rate)	95% CI 0.92 to 1.10 (for 1% increase in truancy rate)		
School categorisation – based on value-added measure – calculated by regressing attainment and truancy rates on five indicators of student profile (% white pupils, % female, mean Townsend score of pupils' ward of residence, % owner-occupiers, % free school meals); school residuals then used in a principal components analysis to create measure of value added accounted for 77% of variance. Authoritative – 1 SD above average; laissez-faire – 1 SD below average; indeterminate – all others	Authoritative OR=0.83; laissez-faire: OR=1.13	Authoritative 95% CI 0.73 to 0.95; laissez-faire 95% CI 1.04 to 1.24	Year 7: 14%; year 9: 13.5%; year 11: 4.9%	The description of the categories of the value-added school culture is debatable because they reflect a measure of value added based on additional effect on school achievement and truancy, controlling for socioeconomic intake, and are therefore assumed to work only as proxies for the school culture
School achievement: five grade A–C GCSEs measured as above	OR=0.94	95% CI 0.88 to 1.01		Only 61% school response rate; additionally, 1476 pupils left out because they went to middle schools where they do not take GCSEs. Pupils providing inconsistent or missing data were excluded. Models contained random intercept for school identifier and random effects for school grade (except the model restricted to grade 7 pupils)
School truancy: measured as above	OR=0.96	95% CI 0.85 to 1.08		
Value added measure – calculated as above (accounted for two-thirds of the variance)	OR=0.87	95% CI 0.78 to 0.95	9.4%	
School achievement: five grade A–C GCSEs measured as above	OR=0.93	95% CI 0.89 to 0.97		
School truancy: measured as above	OR=0.99	95% CI 0.92 to 1.07		
Value-added as measured above (accounted for two-thirds of the variance)	OR=0.91	95% CI 0.85 to 0.96	Grade 7: 9.1%; Grade 9: 11.2%; Grade 11: 0%	
Interaction between value-added school culture measure and school grade	Grade 7: OR=0.57; grade 9: OR=0.65; grade 11: OR=0.96	$\chi^2(2)=10.41$, $p<0.01$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Regular illicit drug use – cannabis, ecstasy, amphetamines, lysergic acid diethylamide (LSD), cocaine, magic mushrooms or heroin (behavioural, dichotomous, self-report)			Grade, sex, ethnicity, SES (housing tenure, free school meal status, Townsend index) whether or not drinks with parents	
	Regular illicit drug use – cannabis, ecstasy, amphetamines, LSD, cocaine, magic mushrooms or heroin (behavioural, dichotomous, self-report)				
Bradshaw <i>et al.</i> 2009 ⁸⁹	Frequent victimisation – respondents were classified as frequently victimised if they reported being bullied two or more times within the past month (dichotomous, experiential, self-report, anonymous)	Elementary schools: 32.18; middle schools: 27.63	Elementary schools: 0.0190; middle schools: 0.0057	Sex, ethnicity, grade status (grades 4 and 6 coded as low status because they are entry grades to schools) school urbanicity, percentage of students receiving free or reduced-cost meals, school level elementary or middle	Student mobility (% migration into and out of school enrolment during school year)
	Perceptions of safety – whether or not respondent feels safe at school (dichotomous, attitudinal, self-report, anonymous)	Elementary schools: 83.73; middle schools: 73.16	Elementary schools: 0.0566; middle schools: 0.0495	Sex, ethnicity, grade status (grades 4 and 6 coded as low status because they are entry grades to schools), school urbanicity, percentage of students receiving free or reduced-cost meals, school level elementary or middle	Student mobility (% migration into and out of school enrolment during school year)
Collins <i>et al.</i> 2008 ⁷²	Lifetime marijuana use (dichotomous, behavioural, self-report)	17 (0.37)	0.04	Sex, race, age, lives in urban/rural/rural farm residence, living with both natural parents, family conflict, parental attitudes favourable to drug use, family attachment, % 0–17 year olds living in poverty in county	Attitudes favourable to drug use, friends' drug use, academic failure, perceived availability of drugs, perceived risk of drug use, commitment to school, school problem scale (mean of perception items measured at pupil level) school safety (mean of items measured at pupil level) self-reported delinquent behaviour

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School achievement: measured as above	OR=0.97	95% CI 0.92 to 1.03		
School truancy: measured as above	OR=1.06	95% CI 0.96 to 1.17		
Value added as measured above (accounted for two-thirds of the variance)	OR=0.90	95% CI 0.82 to 0.98	Grade 7: 0%; Grade 9: 2.5%; Grade 11: 0.2%	
Interaction between value-added school culture measure and school grade	Grade 7: OR=0.78; grade 9: OR=0.75; grade 11: OR=0.79	$\chi^2(2)=0.069$, $p>0.05$		
Student-teacher ratio	Elementary schools: coefficient=0.014, OR=1.01; middle schools: coefficient=0.134, OR=1.14	Elementary schools: $p>0.05$, SE=0.011; middle schools: $p<0.05$, SE=0.035	NR	Final sample contained data on 74% of students in targeted grades
School suspensions (number of suspensions divided by school enrolment)	Elementary schools: coefficient=-0.001, OR=1.00; middle schools: coefficient=0.006, OR=1.01	Elementary schools: $p>0.05$, SE=0.012; middle schools: $p>0.05$, SE=0.003		
Student-teacher ratio	Elementary schools: coefficient=-0.042, OR=0.96; middle schools: coefficient=0.038, OR=1.04	Elementary schools: $p<0.05$, SE=0.012; middle schools: $p>0.05$, SE=0.077	NR	
School suspensions (number of suspensions divided by school enrolment)	Elementary schools: coefficient=0.005, OR=1.01; middle schools: coefficient=-0.020, OR=0.98	Elementary schools: $p>0.05$, SE=0.015; middle schools: $p<0.05$, SE=0.006		
School size	OR=1.00	$p<0.01$	NR	Information related to the covariate 'attitudes favourable to drug use' is not included in the model reporting

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Past year marijuana use (dichotomous, behavioural, self-report)	13 (0.33)	0.04	As above	As above
	Past 30-day marijuana use (dichotomous, behavioural, self-report)	8 (0.27)	0.05	As above	As above
	Lifetime inhalant (solvent) use (dichotomous, behavioural, self-report)	17 (0.37)	0.03	As above	As above
	Past year inhalant (solvent) use (dichotomous, behavioural, self-report)	9 (0.28)	0.04	As above	As above
	Past 30-day inhalant (solvent) use (dichotomous, behavioural, self-report)	5 (0.22)	0.05	As above	As above
	Lifetime other drugs use (dichotomous, behavioural, self-report)	17 (0.38)	0.03	As above	As above
	Past year other drugs use (dichotomous, behavioural, self-report)	12 (0.33)	0.02	As above	As above
	Past 30-day other drugs use (dichotomous, behavioural, self-report)	8 (0.26)	0.03	As above	As above
Comeau 2005 ⁸⁵	Substance use (based on two items: How many cigarettes have you smoked, even a puff, in the last month? How many drinks of alcohol have you had in the last month?)	All students: 1.88 (2.22); males: 1.97 (2.42); females: 1.80 (2.04) (not % prevalence – interval index)	0.027	School SES measured through % receiving free or reduced-cost meals	Deviant peer affiliations
	As above	As above	As above	As above	As above
	As above	As above	As above	As above	As above

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
As above	OR=1.00	$p < 0.01$	NR	
As above	OR=1.00	$p > 0.05$	NR	
As above	OR=1.00	$p < 0.05$	NR	
As above	OR=1.00	$p > 0.05$	NR	
As above	OR=1.00	$p > 0.05$	NR	
As above	OR=1.00	$p < 0.05$	NR	
As above	OR=1.00	$p > 0.05$	NR	
As above	OR=1.00	$p > 0.05$	NR	
Positive emotional environment (principal component analysis component based on three items)	Coefficient=-0.09	SE=0.06, $p > 0.1$	NR	Data were collected on approximately 70% of students. The modelling approach in this doctoral dissertation seems confused and is very poorly reported. It is very difficult to tell exactly what the model specifications were
Observed student deviance	Coefficient=0.09	SE=0.05, $p > 0.1$	NR	
Staff monitoring (principal component analysis component based on two items)	Coefficient=0.08	SE=0.07, $p > 0.1$	NR	

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Cradock <i>et al.</i> 2007 ¹³³	Physical activity during school hours measured using TriTrac-R3D accelerometers (behavioural, monitored, interval)	NR	NR	Time of day, day of week, sex, ethnicity, age, BMI	
	Physical activity during school hours measured using TriTrac-R3D accelerometers	NR	NR	As above	
	Physical activity during school hours measured using TriTrac-R3D accelerometers	NR	NR	As above	
	Physical activity during school hours measured using TriTrac-R3D accelerometers	NR	NR	As above	
	Physical activity during school hours measured using TriTrac-R3D accelerometers	NR	NR	As above	
	Physical activity during school hours measured using TriTrac-R3D accelerometers	NR	NR	As above	
Evans-Whipp <i>et al.</i> 2010 ¹³¹	Smoking in past 30 days (behavioural, self-report, dichotomy)	Washington: 9.5; Victoria: 18.5	NR	State, sex, age, family SES	
	Daily smoking (behavioural, self-report, dichotomy)	Washington: 2.3; Victoria: 7.1	NR	As above	

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Days per week of PE	Coefficient=0.135	SE=0.0457, $p<0.01$	NR	Environmental covariates were measured retrospectively in 2004–5
Campus area per student	Coefficient=0.2244	SE=0.05832, $p<0.001$	NR	
Days per week of PE	Coefficient=0.08234	SE=0.04639, $p>0.05$	NR	
Playground area per student	Coefficient=0.347	SE=0.1474, $p<0.05$	NR	
Days per week of PE	Coefficient=0.1257	SE=0.04548, $p<0.01$	NR	
School building area per student	Coefficient=2.1302	SE=0.9235, $p<0.05$	NR	
Days per week of PE	Coefficient=0.0904	SE=0.04934, $p>0.05$	NR	
Comprehensive smoking ban	OR=0.86	95% CI 0.59 to 1.25, $p=0.42$	NR	
Harsh penalties for smoking	OR=0.99	95% CI 0.73 to 1.35, $p=0.95$		
Remedial penalties for smoking	OR=1.15	95% CI 0.88 to 1.51, $p=0.30$		
Strict enforcement of smoking policy	OR=0.78	95% CI 0.57 to 1.05, $p=0.10$		
Smoking abstinence policy	OR=0.93	95% CI 0.83 to 1.04, $p=0.20$		
Smoking harm minimisation policy	OR=1.09	95% CI 0.99 to 1.21, $p=0.09$		
Comprehensive smoking ban	OR=0.95	95% CI 0.53 to 1.69, $p=0.85$	NR	
Harsh penalties for smoking	OR=1.02	95% CI 0.62 to 1.67, $p=0.95$		
Remedial penalties for smoking	OR=1.10	95% CI 0.72 to 1.69, $p=0.66$		
Strict enforcement of smoking policy	OR=0.70	95% CI 0.44 to 1.12, $p=0.14$		
Smoking abstinence policy	OR=0.95	95% CI 0.80 to 1.13, $p=0.56$		
Smoking harm minimisation policy	OR=1.01	95% CI 0.85 to 1.20, $p=0.89$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Perceptions of many students smoking on school grounds (behavioural, self-report, dichotomy)	Washington: 34.9; Victoria: 63.8	NR	As above	
Gastic 2005 ⁷⁶	Student attitude about school safety based on agreement with statement 'I feel safe at my school' (attitudinal, self-reported, interval)	3.79 (0.03) (not % prevalence: 1 = strongly disagree, 5 = strongly agree)	0.11	Sex, ethnicity, sexual minority (gay, bisexual), grade, annual household income, school urbanicity	Self-reported student delinquency in school, student delinquency reported by school principal, average self-reported student delinquency by school, school connectedness, parental attitude about school safety, friends' average attitude about school safety, attitude misalignment with level of student delinquency at school, self-reported student delinquency in general, self-reported victimisation by violence
Gladden 2005 ⁸⁶	Student perception of feeling unsafe in the classroom (attitudinal, self-report, index, dichotomous)	10		Sex, social status of community where student lives, free/reduced-cost lunch eligibility, whether or not student was old for grade, ethnicity, whether or not student was new to school, whether or not student had attended two or more schools during base-year, rate of violent crime in students' community, student perception of social resources/cohesion in community, racial composition of school, level of violence in community surrounding school, school-level social resources/cohesion in community based on student aggregation, dummy variables to identify missing data bias	Whether or not student was repeating a grade, whether or not school is a magnet school or has a selective enrolment policy, whether or not student has been placed in special education

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments	
Comprehensive smoking ban	OR=1.39	95% CI 0.67 to 2.89, $p=0.38$	NR		
Harsh penalties for smoking	OR=0.72	95% CI 0.42 to 1.22, $p=0.22$			
Remedial penalties for smoking	OR=1.16	95% CI 0.70 to 1.91, $p=0.56$			
Strict enforcement of smoking policy	OR=0.45	95% CI 0.25 to 0.82, $p=0.009$			
Smoking abstinence policy	OR=0.89	95% CI 0.71 to 1.11, $p=0.31$			
Smoking harm minimisation policy	OR=1.18	95% CI 0.97 to 1.43, $p=0.10$			
School size (enrolment)	Standardised coefficient=-0.02	$p<0.01$	0.69	Many other models were conducted, each with a single interaction added. Model 5 was used to provide the information presented here	
Average class size	Standardised coefficient=-0.00	$p>0.05$			
School type (public)	Standardised coefficient=-0.21	$p<0.05$			
Delinquency at school (reported by principal)	Standardised coefficient=-0.13	$p>0.05$			
Code of conduct	Standardised coefficient=-0.01	$p>0.05$			
Security guard	Standardised coefficient=-0.06	$p>0.05$			
School patrols	Standardised coefficient=-0.06	$p>0.05$			
Metal detector	Standardised coefficient=-0.19	$p<0.05$			
Size of the school (reference medium-sized school)	Small: coefficient=-0.35; large: coefficient=0.01	Small: $p<0.10$; large: $p>0.10$	0.283-0.133/ 0.283=0.530		The reporting of the modelling in this PhD dissertation is confusing and does not inspire confidence

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Student perception of feeling unsafe in school hallway/bathroom (attitudinal, self-report, index, dichotomous)	18		As above	As above
	Ever involved in a fight at school during the school year (behavioural, self-report, dichotomous)	25		As above	As above
Grunseit <i>et al.</i> 2008 ⁹⁷	Whether or not student has physically attacked another student in the last 12 months to hurt them, at school or on their way to/from school (behavioural, self-report, dichotomous, confidential)	Overall: 43.7; male: 55.5; female: 31.6	0.064	Sex, whether student lives with both/one/neither parent, mother's age, parenting style (punitive), parenting style (supervisory), problems with family, student impulsiveness	Self-reported problems reading/writing, student opinions of racism of students at school, student perceptions of formal presentation of school rules, student perceptions of justice in response to breaking school rules, student perceptions of teaching style, student perceptions of teacher ability to keep control of class, student perceptions of school response to racism, student perceptions of school response to bullying
Haley 2007 ⁷⁷	Heavy drinking (five or more drinks on one occasion, three to 12 times in the last 12 months) (behavioural, self-report, dichotomous)	17.2	0.1356	Same-sex attraction, sex, ethnicity, age, being 'born again', school urbanicity,	Learning about drinking in class at school, learning about handling conflict in class at school, student has repeated a grade, student has been suspended, grade point average, student has trouble getting along with teachers, student has trouble getting along with students, school connectedness, not being treated fairly by teachers, not feeling safe at school, feeling that teachers do not care
	Drinking-related problems (two or more alcohol-related incidents over the last 12 months) (behavioural/experiential, self-report, dichotomous)	22.7	0.1297	As above	As above

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Size of the school (reference medium-sized school)	Small: coefficient = -0.28; large: coefficient = 0.08	Small: $p < 0.05$; large: $p > 0.10$	0.287–0.108/ 0.287 = 0.624	
Size of the school (reference medium-sized school)	Small: coefficient = -0.26; large: coefficient = -0.00	Small: $p < 0.05$; large: $p > 0.10$	0.150–0.094/ 0.150 = 0.373	
>25% of teachers with <5 years' experience	OR = 1.56	$p = 0.02$	NR	Random slopes on all variables were attempted but found not to vary by school
Strict school policies on punishment	OR = 1.265295	$p = 0.062$	NR	Data weighted at the student level; school-level weights were not available. Missing data bias was identified but ignored. Cases with missing data were deleted listwise. Heavy drinking model OR for size of school is not reported, despite coefficient being previously reported higher up in same table (pp.143–4) – the OR has been calculated from the coefficient by the data extractor. Drinking-related problems model OR for school type is not reported, despite coefficient being previously reported higher up in same table (pp.149–50) – the OR has been calculated from the coefficient by the data extractor
School type (public)	OR = 0.571125	$p = 0.023$		
Size of school (1 = large, >1000)	OR = 1.362 (see reviewer's comments)	$p = 0.026$		
High percentage of student dropouts in school (dichotomous – top 20% of schools by proportion of dropouts)	OR = 0.970056	$p = 0.071$		
School attendance (dichotomous – schools with attendance of $\geq 95\%$)	OR = 0.973533	$p = 0.809$		
Strict school policies on punishment	OR = 1.132123	$p = 0.213$	NR	
School type public	OR = 0.608 (see reviewer's comments)	$p = 0.007$		
Size of school (1 = large, >1000)	OR = 1.187373	$p = 0.190$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Fighting in school (behavioural, self-report, dichotomous)	25.1	0.038	As above	As above
Haug <i>et al.</i> 2010 ⁸⁷	Daily physically active during recess (behavioural, self-report, dichotomous, confidential)	Boys: 41.5; girls: 32.6	NR	SES, sex	Index of student interest in school physical activity
Henry <i>et al.</i> 2009 ⁶⁷	Alcohol use (mean of four standardised items – index) (behavioural, self-report, ratio, confidential)	Within school: 0.00 (0.87); between school: 0.02 (0.20) (not % prevalence: index created from standardised items)	0.05	Sex, ethnicity, percentage in school receiving free or reduced-cost school lunch, average age of students surveyed by school, percentage of students who are white by school, year that the survey took place, rurality of the community, predominant ethnicity of the community	(Although these explanators form the focus of interest in the paper, the authors' first model does not contain them. The results for this first model are reported here – thus, do not contain these inappropriate covariates) School–student bonding, school–friend bonding, student behaviour
Ishibashi 2009 ⁸³	7-day recall of MVPA recoded to hours per week (behavioural, self-report, interval, confidential)	17.3 (11.7) (not % prevalence)	0.019	Average income of school location by school census tract, age, sex, questionnaire module answered	Student-reported average non-participation in school PE by school

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
High percentage of student dropouts in school (dichotomous – top 20% of schools by proportion of dropouts)	OR=0.986176	$p=0.463$		
School attendance (dichotomous – schools with attendance of $\geq 95\%$)	OR=1.048488	$p=0.660$		
Strict school policies on punishment	OR=1.004920	$p=0.956$	NR	
School type (public)	OR=0.938901	$p=0.534$		
Size of school (1=large, >1000)	OR=1.029275	$p=0.719$		
High percentage of student dropouts in school (dichotomous – top 20% of schools by proportion of dropouts)	OR=1.009981	$p=0.336$		
School attendance (dichotomous – schools with attendance of $\geq 95\%$)	OR=0.974921	$p=0.730$		
Availability of facilities for physical activity (environment index)	Coefficient=1.24	$p<0.001$	NR	Original sample included 115 schools with total of 2754 students. Missing data resulted in sample as described here. Some missing data at school level was multiply imputed. Research poorly reported. No explanation of what the policy index variable is and little description of how the modelling was undertaken
School physical activity in school day (policy index) (see reviewer notes)	Coefficient=0.62	$p<0.001$		
Interaction of policy index with environment index	Coefficient=-0.67	$p=0.42$		
Interaction of index of student interest with policy index	Coefficient=1.09	$p=0.22$		
Number of students in school size	Coefficient=0.00	SE=0.00, $p>0.05$	27%	This paper has a very robust modelling design, which makes use of latent variable approaches
Pupil-teacher ratio	Coefficient=0.00	SE=0.00, $p>0.05$		
Whether or not eighth graders in same school as high school students	Coefficient=0.01	SE=0.03, $p>0.05$		
School size	Coefficient=0.1311	$p=0.0210$	NR	

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Johnson and Hoffman 2000 ⁷¹	Beginning to smoke daily between waves 1 and 2 (behavioural, self-report, dichotomous)	Eighth-grade panel: 15.3 (0.6); tenth-grade panel: 12.9 (0.6)	NR	Sex, ethnicity, student works ≥ 10 hours in most recent week, two biological parents at home, parental education, family income, school region (west), school urbanicity, percentage of ethnic minority students in school	Student self-esteem, dropout rate, college plans of students, grade point average, school misconduct, positive school attitude, parental support regarding school (post entry to school), competitive school climate, negative peer associations
Kairouz and Adlaf 2003 ⁷⁹	Five or more alcoholic drinks on a single occasion at least once in the last 4 weeks (behavioural, self-report, dichotomous)	39	0.11	Age, sex, home language, mean age of students in school, ethnolinguistic mix of school, sex composition of school, school average age	Years since first drinking, perceived harmfulness of drinking, disapproval of drinking, permissiveness of school peer-group drinking culture (aggregate of student self-reports)
Khoury-Kassabri et al. 2009 ⁷⁸	Violence towards peers students (behavioural, self-report, interval, confidential)	Threatened to hurt or hit another student: 21.47; kicked or punched another student: 22.76; used a chair, rock or other object to hurt a student: 7.53	0.1223	Sex, grade, ethnic affiliation of school (Jewish/Arab), SES of students' families by school income and education	Student perceptions of school policy, student perceptions of teacher support, student perceptions of participation in school, student perceptions of how teachers deal with violence (principal and homeroom teacher)
	Violence towards teachers (behavioural, self-report, interval, confidential)	Cursed or humiliated a teacher: 14; threatened to hurt a teacher: 3.87; shoved or hit a teacher: 3.68; bit or used a chair to hurt a teacher: 2.13; destroyed personal belongings of a teacher: 6.93	0.14	As above	As above
Kim 1999 ⁸¹	Monthly alcohol use (behavioural, self-report)	NR	0.04081	Locus of control, parental concern	Importance of studying with friends, number of parents who school staff met with this year, prevalence of alcohol use at the school, number of close friends who dropped out of school

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Catholic school	Eighth-grade panel: OR= 1.46; tenth-grade panel: OR= 1.13	Eighth-grade panel: $p < 0.01$; tenth-grade panel: $p > 0.05$	The information available does not match the data requirement	Sample restricted to students who were non-smokers in wave 1 – thus, samples overlap to some degree but are independent. Response rate 92% for responding to both waves. Multiple imputations used to fill in missing values. The modelling was exploratory and so all first-order interactions with sex and ethnicity were attempted. Only the one with the determinant reported here (Hispanic students by students per full-time teacher) remained in the model
Small school size (< 600 students)	Eighth-grade panel: OR=0.91; tenth-grade panel: OR=0.92	Eighth-grade panel: $p > 0.05$; tenth-grade panel: $p > 0.05$		
Students per full-time teacher	Eighth-grade panel: OR=0.93; tenth-grade panel: OR= 1.00	Eighth-grade panel: $p > 0.05$; tenth-grade panel: $p > 0.05$		
Interaction of Hispanic students with students per full-time teacher	Eighth-grade panel: OR=0.81; tenth-grade panel: OR=0.80	Eighth-grade panel: $p < 0.05$; tenth-grade panel: $p < 0.05$		
School size	Coefficient=0.00	$p > 0.05$	91%	
School size	Coefficient=-0.032	$p < 0.01$	65.29%	The school sample was weighted to represent the Israeli student body
Class size	Coefficient=-0.036	$p < 0.05$		
School size	Coefficient=-0.016	$p > 0.05$	45.73%	
Class size	Coefficient=-0.032	$p < 0.05$		
School policy for promoting parental support	Coefficient=-0.032498	$p = 0.075$	NR	The reporting of the analysis in this PhD thesis is incomplete: the details of the sample used in the analysis are missing; the exact definition of the outcomes modelled is missing; the baseline prevalence of the outcome is missing; the model reporting is not clear (seems to be missing main effects when interactions are reported?)

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Daily cigarette use (behavioural, self-report)	NR	0.02144	Locus of control, parental concern, parental divorce	Importance of studying with friends, number of parents who school staff met with this year, number of close friends who dropped out of school
	Monthly marijuana use (behavioural, self-report)	NR	0.01445	Locus of control, parental concern, attending religious activities, parental divorce	Number of parents who school staff met with this year, drug use a problem at the school, number of close friends who dropped out of school
Kim 2001 ⁸⁸	Violent delinquency (behavioural, self-report, dichotomous)	0.24 (0.43) (range 0–3) (not % prevalence: based on scale – average of three items)	0.099/ (0.099 + 0.165 + 1 = 0.0783)	Parental attachment, parental control/supervision, living with two parents, impulsiveness, sex, age, ethnicity, neighbourhood population density, neighbourhood population proportion who are non-white, neighbourhood residential mobility, neighbourhood SES, neighbourhood attachment level, neighbourhood proportion of adolescents not enrolled in school, school proportion male, school proportion non-white, school location (urban)	School performance (grade point average), students' average attachment to school, friends' delinquency level
Kumar <i>et al.</i> 2008 ⁶⁸	Twelfth grade cigarette use in the past 30 days (behavioural, self-report, dichotomous)	Twelfth grade: 27 (0.45)	NR	Alcohol and drug use in the neighbourhood surrounding the school, sex, ethnicity, parental education, lives with both parents, school urbanicity, predominant ethnicity of school, school average SES, percentage students from broken homes by school	
			NR	As above	
	Twelfth graders being drunk in the past 30 days (behavioural, self-report, dichotomous)	Twelfth grade: 32 (0.47)	NR	As above	
	Eighth-grade marijuana use in the past 12 months (behavioural, self-report, dichotomous)	Eighth grade: 15 (0.35); twelfth grade: 36 (0.48)	NR	As above	

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School policy for promoting parental support	Not included in final model		NR	
Interaction of school programme for reducing negative impacts of family background with parental divorce	Coefficient = -0.154888	$p = 0.148$		
School policy for promoting parental support	Not included in final model		NR	
School programme for reducing negative impacts of family background	Coefficient = -0.010687	$p = 0.384$		
Interaction of school programme for reducing negative impacts of family background with parental divorce	Coefficient = -0.164456	$p = 0.001$		
Class size	Coefficient = 0.01	SE = 0.008, not significant (see reviewer's comments)	39%	Sample was weighted in analyses. A cross-nested contextual-effects model using an index outcome is also reported but the dichotomous model is preferred for the purposes of the current study. Alpha level for significance tests is not reported; presumably a 5% level was used?
School level of discipline/regulation	Coefficient = 0.27	SE = 0.101, significant (see reviewer's comments)		
Type of school (public/private)	Estimates not reported for this 'control' variable		NR	Analyses are weighted. Many separate models with one determinant in each, by grade, were developed. The model estimates for the determinants of interest to the authors were the only reported model estimates in the paper
School size	Estimates not reported for this 'control' variable			
Neglected physical environment	Twelfth grade: coefficient = -0.149	Twelfth grade: $p > 0.05$	NR	
Attractive physical environment	Twelfth grade: coefficient = 0.061	Twelfth grade: $p > 0.05$	NR	
Attractive physical environment	Eighth grade: coefficient = 0.145	Eighth grade: $p > 0.05$	NR	

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Twelfth grade marijuana use in the past 12 months (behavioural, self-report, dichotomous)	Eighth grade: 15 (0.35); twelfth grade: 36 (0.48)	NR	As above	
	Alcohol use in school in the past year	Eighth grade: 3 (0.18); tenth grade: 8 (0.27); twelfth grade: 9 (0.28)	NR	As above	
	Use of marijuana and other illicit drugs in school in the past year	Tenth grade: 9 (0.28)	NR	As above	
	Use of cigarettes in the past 4 weeks in school	Eighth grade: 5 (0.23); twelfth grade: 12 (0.45)	NR	As above	
Leatherdale <i>et al.</i> 2010 ⁹⁰	Physical activity level [moderately active (MA) or highly active (HA) compared with low active (LA)] (behavioural, self-report, dichotomous)	LA: 16.4; MA: 67.2; HA: 16.4	MA: 0.048; HA: 0.073	Average screen time per day, league/team sports participation outside of school, varsity/team sports participation in school (because student is talented)	Overweight, how many close friends are physically active, intramural/house league sports participation in school (reported by student)

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Total number of unobservable and unsupervised places in and around school	Twelfth grade: coefficient=0.043	Twelfth grade: $p>0.05$	NR	
Attractive physical environment	Eighth grade: coefficient=0.001; tenth grade: coefficient=-0.350; twelfth grade: coefficient=0.350	Eighth grade: $p>0.05$; tenth grade: $p<0.01$; twelfth grade: $p<0.01$	NR	
Total number of unobservable and unsupervised places in and around school	Tenth grade: coefficient=0.142; twelfth grade: coefficient=0.142	Tenth grade: $p<0.01$; twelfth grade: $p<0.01$		
Total number of unobservable and unsupervised places in and around school	Tenth grade: coefficient=0.100	Tenth grade: $p<0.05$	NR	
Attractive physical environment	Eighth grade: coefficient=0.234	Eighth grade: $p>0.05$	NR	
Total number of unobservable and unsupervised places in and around school	Twelfth grade: coefficient=0.158	Twelfth grade: $p>0.05$		
Student access to facilities on and off school grounds during school hours [healthy physical environment (HPE), see reviewer's comments]	Dropped out in preliminary modelling		NR	50% response rate to convenience sample. Contextual interactions (between student- and school-level variables) were investigated but most were not significant. An interaction between community partnerships (overall score) and league/team sports participation outside of school was significant for HA vs LA but was not reported. The 'school-level indicators' were HPE, instruction and programmes (IP), SSE and community partnerships. They were classed into one of three categories: initiation (lowest class), action (middle class) and maintenance (best class). The following school-level indicators dropped out during the preliminary univariate modelling stage: availability of physical activities during inclement weather (HPE), student access to facilities and equipment outside of school hours (HPE), support for active transportation to/from school (HPE), implementation of daily physical activity (IP), time spent per week engaged in physical activity during PE classes (IP), classes taught by a qualified PE specialist (IP), availability and use of intramural/club activities (IP), consistency of

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Lothian 2003 ⁸⁰	Tried at least one cigarette in the last 12 months (behavioural, self-report, dichotomous, confidential)	Male: 41.6; female: 43.6	0.11	Sex, grade, grade squared	Self-esteem, attachment to substance-using peers, student engagement, student engagement by school, ASB, ASB by school, academic achievement, academic achievement squared, academic achievement by school
	Number of cigarettes smoked per day (behavioural, self-report, interval, confidential) (data set reduced to those who had smoked at least once in the previous 12 months, n=4741)	Male: 3.29 (95% CI 3.20 to 3.38); female: 2.99 (95% CI 2.92 to 3.06)	0.09	As above	As above

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
				intramural programming across grade divisions and seasons (IP), availability and use of interschool programmes (IP), consistency of interschool programming across seasons (IP), emphasis placed on maximising participation in physical activity through school programmes (SSE), incorporation of physical activity into other school subjects (SSE), special recognition of students who participate in school physical activities (SSE), formal collection of suggestion from the school community about physical activity at school (SSE), promotion of physical activity programmes and events for students, families and school staff (SSE), presence of written policies/practices for physical activity (SSE)
Interaction of school size with sex	Coefficient: school size: -0.00002; interaction with sex: 0.00003	School size: $p > 0.05$; interaction with sex: $p > 0.05$	NR	Overall response rate 60%. Some pages from this PhD dissertation were missing. The model for the number of cigarettes smoked per day was based on a subset (less than half) of the sample, as it was restricted to those who had smoked at least once in the last 12 months
Interaction of school size with grade	Coefficient: school size: -0.0002; interaction with grade: 0.00001	School size: $p > 0.05$; interaction with grade: $p > 0.05$		
Interaction of school size with self-esteem	Coefficient: school size: -0.0002; interaction with self-esteem: 0.00002	School size: $p > 0.05$; interaction with self-esteem: $p > 0.05$		
Interaction of school size with attachment to substance-using peers	Coefficient: school size: 0.0001; interaction with attachment: -0.00005	School size: $p > 0.05$; interaction with attachment: $p > 0.05$		
Interaction of school size with sex	Coefficient: school size: 0.00006; interaction with sex: -0.0002	School size: $p > 0.05$; interaction with sex: $p > 0.05$	NR	
Interaction of school size with grade	Coefficient: school size: 0.0003; interaction with grade: -0.0001	School size: $p > 0.05$; interaction with grade: $p < 0.05$		
Interaction of school size with self-esteem	Coefficient: school size: -0.0001; interaction with self-esteem: 0.00002	School size: $p > 0.05$; interaction with self-esteem: $p > 0.05$		
Interaction of school size with attachment to substance-using peers	Coefficient: school size: 0.00005; interaction with attachment: -0.00005	School size: $p > 0.05$; interaction with attachment: $p > 0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Ma 2000 ¹³⁵	Physical health (experiential, self-report scale)	4.93 (4.20) (not % prevalence)	0.02	Sex, SES, whether or not an ethnic native, number of parents at home, number of siblings, school mean SES	General self-esteem, maths/science achievement, reading/writing ability, school disciplinary climate, school academic press (teacher expectations), parental involvement
	Mental health (experiential, self-report, scale)	3.49 (2.69) (not % prevalence)	0.02	As above	As above
	Healthy food eating (behavioural, self-report, scale)	8.15 (5.00) (not % prevalence)	0.17	As above	As above
	Popular (unhealthy) food eating (behavioural, self-report, scale)	2.46 (2.98) (not % prevalence)	0.12	As above	As above
	Sum of exercise occasions (behavioural, self-report, interval)	4.16 (3.17) (not % prevalence)	0.29	As above	As above
Maes and Lievens 1999 ¹⁴⁰	Smoking one or more cigarettes weekly	NR	NR	Mother's smoking, father's smoking, integration with family, student health, school grade	Integration with friends, school results, truancy, repeating classes, attitude towards school, relationship with teachers
	Alcohol use	NR	NR	Sex, student health, school grade	Integration with friends, truancy, repeating classes, attitude towards school, relationship with teachers
	Tooth brushing more than once a day			Year, sex, integration with family	Type of education, attitude to school

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School size	Dropped out		NR	
School size	Dropped out		NR	
School size	Coefficient = -0.01	$p < 0.05$	NR	
School size	Coefficient = -0.01	$p < 0.001$	NR	
School size	Coefficient = -0.01	$p < 0.05$	NR	
High teacher workload	OR = 1.23	$p < 0.05$	Null model between-school variance = 0.328; individual model = 0.105; individual and school model = 0.061	Other behaviours reported as not varying between schools and not explored further: healthy nutrition behaviour, unhealthy nutrition behaviour, physical activity, and medicine use. Other school-level variables examined and reported as not producing statistically significant change in the model fit (log-likelihood test): school size, type of education offered, class size, age structure of teachers, experience of administrator, policy on communication with pupils about performance, policy on pupil evaluation, collaboration between teachers, teacher collegiality, teacher work satisfaction, director-teacher relationship, pupil-teacher relationship, social skills evaluation, evaluation of pupil behaviour, functioning of school committees, evaluation of school policies, healthy school status, drug policy, snack selling at school, possible to eat lunch outside schools
Policy on rules for pupils (not clear on what)	OR = 0.83	$p < 0.05$		
Policy on rules for pupils (not clear on what)	OR = 0.87	$p < 0.05$	Null model between-school variance = 0.435; individual model = 0.100; individual and school model = 0.035	
Female school administrator	OR = 0.66	$p < 0.05$		
Sex ratio teachers male > female	OR = 0.70	$p < 0.05$	Null model between-school variance = 0.215; individual model = 0.073; individual and school model = 0.033	
Female school administrator	OR = 1.39	$p < 0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Malikaew <i>et al.</i> 2003 ¹⁴²	Traumatic dental injury (physiological, dentist examination, dichotomous)	35	NR	Age, marital status of parent, employment status of parent, educational status of parent, family income	Social environment cluster (based on cluster analysis in which main differences were level of supervision by staff, safety topics in school curriculum, participation of parents in school meetings, community activities in school, recorded violence, recorded absenteeism, recorded punishment rates)
				As above	As above
				Age, employment status of parent, educational status of parent, family income	Social environment cluster
				As above	As above
Markham <i>et al.</i> 2008 ⁶⁵	Smoking at least one cigarette per week (behavioural, self-report, dichotomous)	NR	NR	Intervention or control group (data were from a trial that showed no effect of intervention on smoking), whether or not pupil was an occasional smoker at baseline, whether or not pupil was of regular smoking status at baseline, sex, age, ethnicity, student residential district deprivation level, mother's and father's and other relatives' smoking habits	
Moore <i>et al.</i> 2001 ¹³⁶	Weekly smoking	21.8 boys; 29.2 girls	NR	Sex, mother smokes, parents expect too much at school	Alienation
	Daily smoking	18.2 boys; 22.9 girls	NR		

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Physical environment cluster (based on cluster analysis in which main differences were level of school cleanliness, condition of surfaces of playgrounds, crowding)	Boys: OR=1.09, n=1331	95% CI 0.71 to 1.67, $p>0.05$	NR	
Interaction of physical environment cluster with social environment cluster	Boys: OR=0.59, n=1331	95% CI 0.29 to 1.17, $p>0.05$		
Physical environment cluster	Girls: OR=1.02, n=1394	95% CI 0.60 to 1.72, $p>0.05$	NR	
Interaction of physical environment cluster with social environment cluster	Girls: OR=0.61, n=1394	95% CI 0.33 to 1.13, $p>0.05$		
Value-added education measure – calculated as per Aveyard <i>et al.</i> ⁹⁹ but 5-year average GCSE attainment (1995–9) and average school truancy (1995–9), four indicators of social profile (proportion white pupils, proportion females, mean Townsend score of pupils' ward of residence, proportion of mothers who smoked). Value added accounted for 65% of variance High-value added schools – ≥ 1 SD above average; low value-added schools – ≤ 1 SD below average; indeterminate value-added schools – all others	1 SD increase in value-added score: first follow-up: OR=0.85; second follow-up: OR=0.80	First follow-up: 95% CI 0.73 to 0.99; second follow-up: 95% CI 0.71 to 0.91	NR	58.4% of sampled schools participated. Attrition and missingness: 92% of year 9 pupils at baseline, 89.1% at first follow-up, 84.6% at second follow-up
Interaction of value-added education with whether or not pupil was of regular smoking status at baseline	Not significant	First follow-up: $\chi^2=0.002$, $p=0.96$; second follow-up: $\chi^2=0.36$, $p=0.55$		
Weak school policy	OR=2.34	95% CI 1.08 to 5.08	0.249 (SE=0.102)	Several models presented, one of which inappropriately adjusts for best friend smokes, but models reported do not
Medium school policy	OR=1.56	95% CI 0.84 to 2.80		
Weak school policy	OR=3.52	95% CI 1.55 to 7.97	0.202 (SE=0.101)	
Medium school policy	OR=2.19	95% CI 1.11 to 4.35		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Monshouwer <i>et al.</i> 2007 ⁶⁹	Episodic heavy drinking (five or more drinks on one occasion during the last 4 weeks) (behavioural, self-report, dichotomous, confidential)	39.7	0.085	Sex, age, alcohol permitted at home by parents, school ethnic composition	
Murnaghan <i>et al.</i> 2007 ⁸⁴	Occasional smoking (less than weekly) vs never smoking (behavioural, self-report, dichotomous, confidential)	52.6 never smoked, 19.9 occasional smokers, 27.5 regular smokers	NR	Sex, wave of data collection, school location	Student perception of school rules on smoking, student perception of punishments for breaking school rules on smoking, friends' smoking behaviour
	Regular smoking (smoking every week) vs occasional smoking (less than weekly) (behavioural, self-report, dichotomous, confidential)		NR	As above	As above
Pinilla <i>et al.</i> 2002 ¹⁴¹	Daily smoking	6.3	NR	Sex, smokers at home	Interest in school, alcohol consumption (occasionally/regularly), best friend smokes
Piontek <i>et al.</i> 2008 ¹³²	Smoking on any of the previous 30 days (behavioural, self-report, dichotomous, confidential)	23.0	NR	Sex, age, weekly pocket money, parents' highest educational qualification, parents' and siblings' smoking status, parental attitude towards students' smoking	

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School type (see reviewer's comments)	Lower general secondary education: OR=0.69; upper general secondary education: OR=0.40; pre-university education: OR=0.26 (reference category pre-vocational education)	Lower general secondary education: 95% CI 0.57 to 0.83, $p < 0.001$; upper general secondary education: 95% CI 0.32 to 0.51, $p < 0.001$; pre-university education: 95% CI 0.20 to 0.33, $p < 0.001$	NR	72% response rate for schools, 7% non-response of selected students. Sample was weighted post stratification to match national statistics. Although school type is classed as a determinant here, the type of school that students go to is determined from tests taken at the end of primary school – thus, the students are likely to be systematically different in different types of school
Alcohol use permitted at school on special occasions	OR=1.11	95% CI 0.92 to 1.34, $p > 0.05$		
Level of sanctioning if caught using alcohol at school	Medium: OR=0.95; high: OR=1.01 (reference category low)	Medium: 95% CI 0.75 to 1.22, $p > 0.05$; high: 95% CI 0.77 to 1.32, $p > 0.05$		
School-based smoking policy implemented	OR=1.06	95% CI 0.67 to 1.68, $p > 0.05$	NR	
Both a school-based smoking prevention programme and school-based smoking policy implemented	OR=0.83	95% CI 0.61 to 1.12, $p > 0.05$		
Interaction of school-based smoking policy implemented with student perception of school rules on smoking	OR=2.69	95% CI 1.20 to 6.03, $p < 0.05$		
School-based smoking prevention programme	OR=1.11	95% CI 0.80 to 1.53, $p > 0.05$	NR	
School-based smoking policy implemented	OR=0.79	95% CI 0.52 to 1.19, $p > 0.05$		
Both a school-based smoking prevention programme and school-based smoking policy implemented	OR=0.88	95% CI 0.65 to 1.19, $p > 0.05$		
School directors' view of compliance with antismoking rules at school	Coefficient=-0.770	SE=0.332	Null model between-school variance=0.433; individual model=0.518; individual and school model=0.477	Other school-level variables examined and reported as having no significant effects: type of school, rate of academic success
School type (Hauptschulen: low level, Realschulen: intermediate level, Gymnasien: high level, Gesamtschulen: intermediate and high level)	Coefficient=-0.17 (see reviewer's comments)	SE=0.12, $p > 0.05$	NR	84% of original data used in analysis ($n=2818$). The school type coefficient seems to indicate a misspecification of the model because only a single coefficient is reported for a categorical variable

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
Rountree and Clayton 1999 ⁷³	Alcohol use (six-category ordinal scale from 1 = never to 6 = daily) (behavioural, self-report, ordinal)	2.45 (1.34) (not % prevalence) (there are serious problems with the use of the outcome variable; see reviewer's comments)	0.103	How often student attends religious services, sex, age, ethnicity, proportion of students living in a rural setting by school, proportion of non-white students by school	Student attachment to school index, peer alcohol use, school-level drug subculture index, school-level criminal subculture index
Tobler <i>et al.</i> 2011 ⁵³	Drank alcohol on one or more occasions in the last 30 days (behavioural, dichotomous, self-report) in grade 8	Sixth grade: 6.57; eighth grade: 22.97	NR	See reviewer's comments	
	Drank five or more alcoholic drinks on one occasion in the last 2 weeks (behavioural, dichotomous, self-report) in grade 8	Sixth grade: 3.40; eighth grade: 8.92	NR	See reviewer's comments	
	Smoked a cigarette in the last month (behavioural, dichotomous, self-report) in grade 8	Sixth grade: 1.29; eighth grade: 9.89	NR	See reviewer's comments	
	Have used marijuana in the last 30 days (behavioural, dichotomous, self-report) in grade 8	Sixth grade: 1.72; eighth grade: 12.10	NR	See reviewer's comments	
	Have been involved in a group fight in the last month (behavioural, dichotomous, self-report) in grade 8	Sixth grade: 24.73; eighth grade: 30.58	NR	See reviewer's comments	

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Complete smoking ban for students at/around school	Coefficient=-0.48, OR=0.62	SE=0.19; 95% CI 0.42 to 0.92, $p < 0.05$		
Extensive smoking ban for adults at/around school	Coefficient=-0.06	SE=0.19, $p > 0.05$		
Student smoking monitored regularly at/around school	Coefficient=0.25	SE=0.22, $p > 0.05$		
School smoking policy implements constructive smoking-related measures sanctions	Coefficient=0.10	SE=0.23, $p > 0.05$		
School size	Coefficient=0.015	SE=0.023, $p > 0.05$	86%	Stratified random 11% subsample ($n=3359$) of original data ($n=29,915$) used, but missing data of 32% resulted in final sample of $n=2295$. This analysis suffers from a serious misspecification whereby it seems that continuous regression approaches have been used on an ordinal outcome variable that should not be approximated as continuous. This approach seems to have also been used with some covariates
Value-added school culture measure – calculated by regressing 3-year averages of school achievement of national norm for reading and mathematics and school truancy on the following 3-year averages of social profile: racial/ethnic composition, proportion of students from low-income families, proportion male, proportion who speak English at home, proportion coming from two-parent homes; school residuals were then used in a principal components analysis to create a single component measure of value added that accounted for 80% of the variance (interval). Value-added category – 1 SD above average; value-attenuated category – 1 SD below average; normative – all others	Value added: OR=0.60; value attenuated: OR=0.96; reference category: normative	Value added: 95% CI 0.42 to 0.88, $p=0.008$; value attenuated: 95% CI 0.63 to 1.47, $p=0.856$	NR	There was a model that contained many appropriate covariates but it also contained an inappropriate covariate and so the base model was preferred. Appropriate covariates in the adjusted model included perceived neighbourhood problems scale, area deprivation index, parental monitoring and communication, sex, race/ethnicity, whether or not in the treatment condition (data are from an intervention RCT). Inappropriate covariates in the adjusted model included student academic problems
	Value added: OR=0.44; value attenuated: OR=0.90; reference category: normative	Value added: 95% CI 0.23 to 0.84, $p=0.013$; value attenuated: 95% CI 0.47 to 1.70, $p=0.744$	NR	
	Value added: OR=0.48; value attenuated: OR=0.72; reference category: normative	Value added: 95% CI 0.26 to 0.86, $p=0.015$; value attenuated: 95% CI 0.37 to 1.39, $p=0.323$	NR	
	Value added: OR=0.29; value attenuated: OR=0.98; reference category: normative	Value added: 95% CI 0.15 to 0.57, $p < 0.001$; value attenuated: 95% CI 0.57 to 1.69, $p=0.954$	NR	
	Value added: OR=0.69; value attenuated: OR=1.02; reference category: normative	Value added: 95% CI 0.50 to 0.96, $p=0.026$; value attenuated: 95% CI 0.68 to 1.51, $p=0.941$	NR	

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
van der Horst et al. 2009 ¹³⁹	Overweight status (physiological, examination, dichotomous)	20.2	NR	Ethnicity, age	
	Unfavourable soft drink consumption (behavioural, self-report, dichotomous)	67.2	NR	As above	
	Unfavourable breakfast consumption (behavioural, self-report, dichotomous)	43.5	NR	As above	
	Unfavourable snack intake (behavioural, self-report, dichotomous)	50.0	NR	As above	
	Unfavourable walking during leisure time (behavioural, self-report, dichotomous)	44.0	NR	As above	
	Unfavourable bicycling during leisure time (behavioural, self-report, dichotomous)	28.5	NR	As above	
	Unfavourable sports playing during leisure time (behavioural, self-report, dichotomous)	19.8	NR	As above	
	Unfavourable active commuting to school (behavioural, self-report, dichotomous)	33.0	NR	As above	
	Unfavourable television viewing (behavioural, self-report, dichotomous)	40.6	NR	As above	
	Unfavourable computer use (behavioural, self-report, dichotomous)	47.1	NR	As above	
	Unfavourable soft drink consumption (behavioural, self-report, dichotomous)		NR	As above	
	Unfavourable soft drink consumption (behavioural, self-report, dichotomous)		NR	As above	

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School type (vocational, reference category is higher-level secondary)	OR= 1.7	95% CI 1.19 to 2.33	NR	72% of original sample remained after missing data and data problems had been removed. Other interactions were investigated (ethnicity with sex and school type); only the significant interactions are reported by the authors
School type (vocational, reference category is higher-level secondary)	OR=2.0	95% CI 1.19 to 3.22	NR	
School type (vocational, reference category is higher-level secondary)	OR= 1.5	95% CI 0.88 to 2.39	NR	
School type (vocational, reference category is higher-level secondary)	OR= 1.5	95% CI 1.16 to 2.00	NR	
School type (vocational, reference category is higher-level secondary)	OR=0.5	95% CI 0.32 to 0.79	NR	
School type (vocational, reference category is higher-level secondary)	OR=0.9	95% CI 0.55 to 1.34	NR	
School type (vocational, reference category is higher-level secondary)	OR=2.5	95% CI 1.37 to 4.43	NR	
School type (vocational, reference category is higher-level secondary)	OR= 1.4	95% CI 0.62 to 3.20	NR	
School type (vocational, reference category is higher-level secondary)	OR= 1.7	95% CI 1.08 to 2.71	NR	
School type (vocational, reference category is higher-level secondary)	OR= 1.4	95% CI 0.90 to 2.14	NR	
Interaction of school type (vocational) with Western ethnicity (subsample analysis)	OR=3.2	95% CI 1.63 to 6.34	NR	
Interaction of school type (vocational) with non-Western ethnicity (subsample analysis)	OR= 1.5	95% CI 0.92 to 2.30	NR	

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Unfavourable television viewing (behavioural, self-report, dichotomous)		NR	As above	
	Unfavourable television viewing (behavioural, self-report, dichotomous)		NR	As above	
	Unfavourable computer use (behavioural, self-report, dichotomous)		NR	As above	
	Unfavourable computer use (behavioural, self-report, dichotomous)		NR	As above	
Way 2003 ⁸²	Number of in-school fights in previous semester (measured in 1990, tenth grade) (behavioural, self-report, ordinal)	0.185 (0.458) (not % prevalence: this is the mean of an ordinal variable, range 0–2, with three categories: none, once or twice, more than twice; see reviewer's comments)	0.049	Sex, ethnicity, lives with both parents, number of siblings, speaks English as a second language, SES, school urbanicity, % African American students in school, % Hispanic students in school, % students in school who receive free school meals, average school SES, % African American teachers in school, % Hispanic teachers in school	Student perception of strictness of school rules, student perception of strictness of school rules squared, student eighth-grade test scores, student eighth-grade misbehaviour, % students who drop out of school, aggregated teacher and student perceptions of school disorder, student eighth-grade in-school fighting, on an academic track, on a vocational track
West <i>et al.</i> 2010 ¹³⁷	Self-esteem (attitudinal, self-report, interval)	NR	NR	Self-esteem at baseline	School concerns and peer concerns – poor experiences of transition between primary and secondary school, pupil disengagement at secondary school, pupil perceptions of poor ethos at secondary school

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Interaction of school type (vocational) with Western ethnicity (subsample analysis)	OR=2.9	95% CI 1.55 to 5.46	NR	
Interaction of school type (vocational) with non-Western ethnicity (subsample analysis)	OR=1.3	95% CI 0.82 to 1.96	NR	
Interaction of school type (vocational) with Western ethnicity (subsample analysis)	OR=2.1	95% CI 1.25 to 3.59	NR	
Interaction of school type (vocational) with non-Western ethnicity (subsample analysis)	OR=1.2	95% CI 0.70 to 1.97	NR	
Punishment severity (weighted composite measure built from multiple items)	Coefficient=0.007	SE=0.006, $p>0.05$	32.70%	Design weights were used in the analysis to account for oversampling of minority populations. Difficult to assess the mean sample size per school used in the analysis because the original sample size reported was 13,371 but that reported in the model extracted was 9279, and the number of schools remaining is not reported. Author has taken an ordinal outcome measure with three categories and used it as a continuous variable with values 0, 1 and 2: serious misspecification of the model and may affect validity of results
School rules (multiple dichotomous items indicating presence of particular rules at school, summed)	Coefficient=0	SE=0.009, $p>0.05$		
Private school	Coefficient=-0.012	SE=0.037, $p>0.05$		
Catholic school	Coefficient=-0.029	SE=0.034, $p>0.05$		
Square root of student-teacher ratio	Coefficient=-0.01	SE=0.012, $p>0.05$		
Log of school size	Coefficient=-0.018	SE=0.013, $p>0.05$		
Primary school is a placing school	Age 13 years: standardised coefficient=0.15; age 15 years: standardised coefficient=0.11	Age 13 years: $p<0.01$; age 15 years: $p>0.05$	NR	
Primary school roll	Age 13 years: standardised coefficient=-0.01; age 15 years: standardised coefficient=-0.01	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school policies (health education, smoking, truancy, bullying)	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school policies (parental involvement)	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Depression (attitudinal, self-report, dichotomous)	NR	NR	Depression at baseline	As above
	Self-esteem at age 18–19 years (attitudinal, self-report, interval)	NR	NR	Self-esteem at baseline	As above
	Psychological distress at age 18–19 years (attitudinal, self-report, dichotomous)	NR	NR	Psychological distress at baseline	As above

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Researcher ratings of secondary school ethos (facilities, teacher-pupil relations, pupil participation in decisions)	Age 13 years: standardised coefficient=0.01; age 15 years: standardised coefficient=0.01	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Secondary school size	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school is a placing school	Age 13 years: OR=0.87; age 15 years: OR=0.75	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school roll (size)	Age 13 years: OR=1.03; age 15 years: OR=1.05	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school policies (health education, smoking, truancy, bullying)	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school policies (parental involvement)	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Researcher ratings of secondary school ethos	Age 13 years: OR=0.98; age 15 years: OR=1.26	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Secondary school size	Not reported because not significant across models	Age 13 years: $p>0.05$; age 15 years: $p>0.05$		
Primary school is a placing school	Not reported because not significant across models	Age 18-19 years: $p>0.05$		
Primary school roll size	Not reported because not significant across models	Age 18-19 years: $p>0.05$		
Primary school policies (health education, smoking, truancy, bullying)	Not reported because not significant across models	Age 18-19 years: $p>0.05$		
Primary school policies (parental involvement)	Not reported because not significant across models	Age 18-19 years: $p>0.05$		
Researcher ratings of secondary school ethos	Age 18-19 years: standardised coefficient=0.08	Age 18-19 years: $p<0.05$		
Secondary school size	Age 18-19 years: standardised coefficient=0.03	Age 18-19 years: $p>0.05$		
Primary school is a placing school	Not reported because not significant across models	Age 18-19 years: $p>0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
West <i>et al.</i> 2004 ¹³⁰	Current regular or occasional smoking (behavioural, dichotomous, self-report)	Age 13 years: 11.9 (7.2); age 15 years: 24.8 (10.9)	Age 13 years: 0.062; age 15 years: 0.037	Previous health behaviours, age, sex, social class, deprivation, religion, family structure, parenting care scale, parenting control scale, student disposable income, parental smoking, parental drinking, previous student engagement with school	School environment rated by students related to physical environment and teaching, student involvement in school, student engagement in school, student perception of teacher/pupil relationships, school ethos measure (mean of aggregate of above measures)
Wiium and Wold 2006 ⁷⁴	Adolescent smokes at least once a week (behavioural, self-report, dichotomous, confidential)	25	0.12	Parental norms on pupils' smoking behaviour, family smoking	Student attitude to smoking, student perceived behavioural control of smoking, student perceptions of teachers' smoking, student perceptions of teachers' smoking by school
	Regular monthly alcohol consumption (behavioural, dichotomous, self-report)	Age 13 years: 31.7 (11.3); age 15 years: 63.2 (12.3)	Age 13 years: 0.044; age 15 years: 0.037	As above	As above
	Ever using illicit drugs (behavioural, dichotomous, self-report)	Age 13 years: 18.4 (7.9); age 15 years: 40.0 (11.6)	Age 13 years: 0.039; age 15 years: 0.031	As above	As above

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
Primary school roll size	Not reported because not significant across models	Age 18–19 years: $p > 0.05$		
Primary school policies (health education, smoking, truancy, bullying)	Not reported because not significant across models	Age 18–19 years: $p > 0.05$		
Primary school policies (parental involvement)	Not reported because not significant across models	Age 18–19 years: $p > 0.05$		
Researcher ratings of secondary school ethos	Age 18–19 years: OR=0.66	Age 18–19 years: $p < 0.05$		
Secondary school size	Age 18–19 years: OR=0.90	Age 18–19 years: $p > 0.05$		
School roll (in quintiles)	Age 13 years: OR=1.18; age 15 years: OR=1.09	Age 13 years: $p < 0.05$; age 15 years: $p > 0.05$	NR	
School enforcement of restrictions on teachers' and pupils' smoking	Low: OR=2.00 (reference category high)	95% CI 0.55 to 7.32	NR	Final sample used in analysis was 68% of original sample size
Interaction of school enforcement of restrictions on teachers' and pupils' smoking with family smoking	Low by family smoking interaction: OR=5.57	95% CI 1.47 to 21.12		
School roll (in quintiles)	Age 13 years: OR=1.03; age 15 years: OR=1.07	Age 13 years: $p > 0.05$; age 15 years: $p > 0.05$		
Rating of school by researchers	Age 13 years: OR=1.11; age 15 years: OR=1.01	Age 13 years: $p < 0.05$; age 15 years: $p > 0.05$		
School denomination (Catholic)	Age 13 years: OR=1.15; age 15 years: OR=0.63	Age 13 years: $p > 0.05$; age 15 years: $p < 0.05$		
School roll (in quintiles)	Age 13 years: OR=1.13; age 15 years: OR=1.05	Age 13 years: $p < 0.05$; age 15 years: $p > 0.05$		
Rating of school by researchers	Age 13 years: OR=1.08; age 15 years: OR=1.09	Age 13 years: $p > 0.05$; age 15 years: $p > 0.05$		
School denomination (Catholic)	Age 13 years: OR=0.69; age 15 years: OR=0.93	Age 13 years: $p > 0.05$; age 15 years: $p > 0.05$		

continued

TABLE 32 Data extraction of multilevel studies (health outcomes) (continued)

Study	Health outcome	Mean prevalence of outcome, % (SD)	ICC	Appropriate covariates	Inappropriate covariates
	Unhealthy diet – based on a ratio of fats to carbohydrates (behavioural, dichotomous, self-report)	Age 13 years: 57.7 (14.9); age 15 years: 61.1 (13.0)	Age 13 years: 0.086; age 15 years: 0.063	As above	As above
Wium and Wold 2009 ⁷⁵	Adolescent smokes daily (behavioural, self-report, dichotomous, confidential)	18.7	0.11	Evening spent with friends, engagement in sport outside of school, parents' smoking behaviour, family structure, parents' disapproval of adolescent smoking, parental academic support, parental emotional support, family wealth	Academic achievement, student liking of school, best friend smoking status

NR, not reported; SE, standard error.

School-level determinants	Effect type = x	Measure of precision = y	Between-school variance explained by school-level variables in model	Comments
School roll (in quintiles)	Age 13 years: OR=0.95; age 15 years: OR= 1.06	Age 13 years: $p > 0.05$; age 15 years: $p > 0.05$		
Rating of school by researchers	Age 13 years: OR= 1.06; age 15 years: OR= 1.30	Age 13 years: $p > 0.05$; age 15 years: $p < 0.05$		
School denomination (Catholic)	Age 13 years: OR=0.89; age 15 years: OR=0.93	Age 13 years: $p > 0.05$; age 15 years: $p > 0.05$		
Number of years teacher smoking policy has been in place	OR=0.93	95% CI 0.74 to 1.16	NR	Final sample used in analysis was 68% of original sample size
Whether or not teachers were involved in development of teacher smoking policy	No: OR= 1.15; don't know: OR= 1.10 (reference category yes)	No: 95% CI 0.64 to 2.07; don't know: 95% CI 0.31 to 3.92		
School enforcement of restrictions on teachers' and pupils' smoking	Low: OR= 1.79 (reference category high)	95% CI 0.93 to 3.42		

Quality appraisal of multilevel studies

Reports were quality assessed using the following criteria: (1) whether or not appropriate covariates (key potential confounders) were adjusted for and (2) whether or not (inappropriate covariates) potential mediators were adjusted for.

TABLE 33 Appropriate covariates adjusted for

	Student sex	Student or staff ethnicity, race or language	Student grade or age	Student sexuality	Student or family religion	Student pocket money, part-time work or income	Student peer associations outside school	Student educational attainment or experience prior to current school
Adams <i>et al.</i> 2009 ¹³⁴	✓	✓	✓					
Akiba 2010 ⁷⁰	✓	✓	✓					
Attar-Schwartz and Khoury-Kassabri 2008 ¹³⁸	✓	✓	✓					
Aveyard <i>et al.</i> 2004 ⁴⁸	✓	✓	✓					
Bisset <i>et al.</i> 2007 ⁶⁶	✓	✓						
Bradshaw <i>et al.</i> 2009 ⁸⁹	✓	✓	✓					
Collins <i>et al.</i> 2008 ⁷²	✓	✓	✓					
Comeau 2005 ⁸⁵								
Cradock <i>et al.</i> 2007 ¹³³	✓	✓	✓					
Evans-Whipp <i>et al.</i> 2010 ¹³¹	✓		✓					
Gastic 2005 ⁷⁶	✓	✓	✓	✓				
Gladden 2005 ⁸⁶	✓	✓	✓					✓
Grunseit <i>et al.</i> 2008 ⁹⁷	✓							
Haley 2007 ⁷⁷	✓	✓	✓	✓	✓			
Haug <i>et al.</i> 2010 ⁸⁷	✓	✓						
Henry <i>et al.</i> 2009 ⁶⁷	✓	✓	✓					
Ishibashi 2009 ⁸³	✓		✓					
Johnson and Hoffman 2000 ⁷¹	✓	✓				✓		
Kairouz and Adlaf 2003 ⁷⁹	✓	✓	✓					
Khoury-Kassabri <i>et al.</i> 2009 ⁷⁸	✓	✓	✓					
Kim 1999 ⁸¹					✓			
Kim 2001 ⁸⁸	✓	✓	✓					
Kumar <i>et al.</i> 2008 ⁶⁸	✓	✓						
Leatherdale <i>et al.</i> 2010 ⁹⁰								

Student health status or behaviours prior to or outside current school	Family SES	Family structure or parental divorce	Parental age	Family functioning, attachment or parenting styles	Local area socioeconomic or health characteristics	Trial intervention arm or study data collection wave	Family health attitudes or behaviours	School level
	✓				✓			
	✓							
	✓				✓		✓	
	✓				✓		✓	
	✓				✓			✓
	✓	✓		✓	✓		✓	
	✓					✓		
	✓				✓			
	✓				✓			
✓		✓	✓	✓				
	✓				✓			
	✓				✓	✓		
	✓	✓			✓	✓		
✓	✓	✓			✓			
	✓							
✓		✓		✓				
✓		✓		✓	✓			
	✓	✓			✓			
✓								

continued

TABLE 33 Appropriate covariates adjusted for (continued)

	Student sex	Student or staff ethnicity, race or language	Student grade or age	Student sexuality	Student or family religion	Student pocket money, part-time work or income	Student peer associations outside school	Student educational attainment or experience prior to current school
Lothian 2003 ⁸⁰	✓	✓	✓					
Ma 2000 ¹³⁵	✓	✓						
Maes and Lievens 1999 ¹⁴⁰			✓					
Malikaew <i>et al.</i> 2003 ¹⁴²			✓					
Markham <i>et al.</i> 2008 ⁶⁵	✓	✓	✓					
Moore <i>et al.</i> 2001 ¹³⁶	✓							
Monshouwer <i>et al.</i> 2007 ⁶⁹	✓	✓	✓					
Murnaghan <i>et al.</i> 2007 ⁸⁴	✓							
Pinilla <i>et al.</i> 2002 ¹⁴¹	✓							
Piontek <i>et al.</i> 2008 ¹³²	✓		✓			✓		
Rountree and Clayton 1999 ⁷³	✓	✓	✓		✓			
Tobler <i>et al.</i> 2011 ⁵³	✓	✓						
van der Horst <i>et al.</i> 2009 ¹³⁹		✓	✓					
Way 2003 ⁸²	✓	✓	✓					
West <i>et al.</i> 2004 ¹³⁰	✓		✓		✓	✓		✓
West <i>et al.</i> 2010 ¹³⁷	✓		✓		✓			✓
Wiium and Wold 2006 ⁷⁴								
Wiium and Wold 2009 ⁷⁵							✓	

Student health status or behaviours prior to or outside current school	Family SES	Family structure or parental divorce	Parental age	Family functioning, attachment or parenting styles	Local area socioeconomic or health characteristics	Trial intervention arm or study data collection wave	Family health attitudes or behaviours	School level
✓								
✓	✓	✓		✓			✓	
	✓	✓						
✓					✓	✓	✓	
				✓			✓	
				✓				
					✓	✓		
					✓		✓	
					✓		✓	
				✓	✓	✓		
	✓	✓			✓			
✓	✓	✓		✓			✓	
✓	✓	✓		✓				
							✓	
✓	✓	✓		✓			✓	

TABLE 34 Covariates inappropriately adjusted for

	School organisation, policies, practices	School climate	Student mobility, repeating, dropout	Student misbehaviour, truancy	Parent attitudes to/ engagement with school
Adams <i>et al.</i> 2009 ¹³⁴					
Akiba 2010 ⁷⁰	✓	✓			
Attar-Schwartz and Khoury-Kassabri 2008 ¹³⁸	✓				
Bradshaw <i>et al.</i> 2009 ⁸⁹			✓		
Collins <i>et al.</i> 2008 ⁷²		✓		✓	
Comeau 2005 ⁸⁵					
Gastic 2005 ⁷⁶				✓	✓
Gladden 2005 ⁸⁶	✓		✓		
Grunseit <i>et al.</i> 2008 ⁹⁷	✓	✓			
Haley 2007 ⁷⁷		✓	✓		
Haug <i>et al.</i> 2010 ⁸⁷					
Ishibashi 2009 ⁸³					
Johnson and Hoffman 2000 ⁷¹		✓	✓	✓	✓
Kairouz and Adlaf 2003 ⁷⁹					
Khoury-Kassabri <i>et al.</i> 2009 ⁷⁸	✓	✓			
Kim 1999 ⁸¹			✓		✓
Kim 2001 ⁸⁸				✓	
Leatherdale <i>et al.</i> 2010 ⁹⁰					
Lothian 2003 ⁸⁰				✓	
Ma 2000 ¹³⁵		✓			✓
Maes and Lievens 1999 ¹⁴⁰			✓	✓	
Malikaew <i>et al.</i> 2003 ¹⁴²		✓			
Moore <i>et al.</i> 2001 ¹³⁶					
Murnaghan <i>et al.</i> 2007 ⁶⁹	✓				
Pinilla <i>et al.</i> 2002 ¹⁴¹					
Rountree and Clayton 1999 ⁷³				✓	
Way 2003 ⁸²	✓	✓			
West <i>et al.</i> 2004 ¹³⁰		✓			
West <i>et al.</i> 2010 ¹³⁷		✓			
Wiium and Wold 2006 ⁷⁴					
Wiium and Wold 2009 ⁷⁵					

School health education, PE, health services	Student educational experiences, attitudes or aspirations	Student educational attainment	Student relationships with staff/other students	Student health attitudes, behaviours or status	No inappropriate covariates adjusted
				✓	
	✓	✓	✓		
	✓				
	✓	✓		✓	
			✓		
	✓			✓	
	✓			✓	
✓		✓			
		✓			
✓		✓		✓	
	✓	✓	✓	✓	
				✓	
	✓			✓	
	✓			✓	
		✓			
	✓		✓		
				✓	
	✓	✓		✓	
				✓	
	✓		✓		
				✓	
	✓	✓		✓	

Research question 5: qualitative studies

TABLE 35 Data extraction of quality studies (study context)

Study	Country	School type; grade/age	Sample	Socioeconomic profile	Urban/rural	Ethnic/racial profile	Research methods
Astor <i>et al.</i> 1999 ¹⁴⁸	USA	High	Students, school staff	Two schools: > 85% of students from economically disadvantaged households; one school: economically diverse; one school: 60% of students from economically disadvantaged households; one school: high socioeconomic profile	Five schools mixture of inner city, semiurban and suburban	One school: 99% African American; one school: 100% African American; one school: 30% Mid-Eastern, 70% European American; one school: 50% African American, 50% European American; one school: > 78% European American, 15% African American	Focus groups, interviews
Astor <i>et al.</i> 2001 ¹⁵³	USA	Middle and elementary	Students	82% of the sample had reduced school meals	Two schools in urban areas	Sample: 51% African American, 40% white, 5% Latino	Interviews, school records, spatial/mapping methods
Brotman 2009 ⁹⁴	USA	High; twelfth grade	Students	NR (school specialises in preparing students for entering training for health professions)	Urban	NR	Focus groups, interviews, observation
Brunson and Miller 2009 ¹⁵⁰	USA	NR (students recruited from alternative school; neighbourhood organisation)	Students (young men considered to be at risk)	Majority of sample: intense racial segregation; poverty, unemployment and female-headed families; and high rates of crime	Urban	African American	Interviews, open-ended questionnaire
Burnett 1998 ¹⁴⁹	South Africa	Standard; 7	Students, adults	Chronic poverty	Rural (township had a total of 4000 residents)	Located in a 'coloured' township in terms of apartheid legislation: mixed heritage	Participant observation, interviews, group discussions, self-concept questionnaire, diaries/essay

continued

TABLE 35 Data extraction of quality studies (study context) (continued)

Study	Country	School type; grade/age	Sample	Socioeconomic profile	Urban/rural	Ethnic/racial profile	Research methods
Cousins 1997 ¹⁶²	USA	High	Students	NR (implicit: area of deprivation)	Urban	Mostly African American	Participant observation
Devine 1995 ¹⁵¹	USA	High	Students	NR (implicit: area of very high deprivation)	Urban	African American	Participant observation
Fletcher <i>et al.</i> 2009 ⁹¹⁻⁹³	UK	Secondary schools; 14–15 years	Students, teachers	(1) Large proportion of middle-class students and a minority of students from low-income families; (2) 50% of students on free school meals; and (3) most students in sample in rental accommodation	Urban	(1) NR; (2) almost all from black/minority backgrounds; and (3) 10 white British, 10 black, 7 dual heritage, 2 white European, 1 Hispanic	Interviews, observation
Gordon and Turner 2001 ⁹⁵	UK	Secondary	Students, teachers	School located in a deprived area	NR	NR	Focus groups, interviews
Haselswerdt and Lenhardt 2003 ¹⁵⁷	USA	Middle and high schools	Students	NR (likely varied across schools studied)	Six urban, suburban and rural districts	NR	Focus groups
Hosie 2007 ¹⁵⁸	UK	NR	Students (young women)	NR	10 local education authorities: urban and rural	Mostly white (across sample); one local education authority predominantly black (African or Caribbean)	Interviews
Keddie 2009 ¹⁵⁹	Australia	Secondary	Students (girls only), teachers	Middle class	NR	Anglo Saxon/white	Focus groups and interviews

Study	Country	School type; grade/age	Sample	Socioeconomic profile	Urban/ rural	Ethnic/racial profile	Research methods
Lundblad <i>et al.</i> 2010 ¹⁶⁰	Sweden	Elementary	Students	NR	NR five schools	NR	Interviews
Marsiglia <i>et al.</i> 2002 ⁹⁶	USA	Middle	Students (mostly young women)	> 50% of sample on free school lunch	Urban	Latina/o (most Mexican or other Hispanic/mixed Hispanic)	Interviews
Pike and Colquhoun 2009 ¹⁶³	UK	Primary; 5–11 years	Students, teachers, lunchtime supervisors	High socioeconomic deprivation	NR	NR	Interviews, focus groups, observations
Plano Clark <i>et al.</i> 2002 ¹⁵⁵	UK	High	Students	NR	Two urban high schools, one suburban, one rural	NR	Focus groups (students as co- researchers)
Waldron 2005 ¹⁶¹	USA	High	Students	(1) Over half from middle class; and (2) half poor or working class	Urban	(1) 83.6% white; and (2) 98.8% African American	Interviews, observation
Wills <i>et al.</i> 2005 ¹⁵⁶	UK	Secondary	Students	Most deprived postcode sectors	NR	NR	Interviews
Wilson- Simmons <i>et al.</i> 2006 ¹⁵²	USA	Middle; grades 6–8	Students, staff	City has high violent crime; three-quarters in rental housing	Urban	90% African American	Focus groups

NR, not reported.

TABLE 36 Quality assessment of qualitative studies

Study	Were steps taken to increase rigour in sampling?	Were steps taken to increase rigour in data collected?	Were steps taken to increase rigour in analysis of data?	Were findings grounded in data?	Findings in terms of breadth	Findings in terms of depth	To what extent are the perspectives of young people privileged?	Overall reliability/trustworthiness
Astor <i>et al.</i> 1999 ¹⁴⁸	+	++	++	+++	+++	+++	+++	High
Astor <i>et al.</i> 2001 ¹⁵³	+	++	++	+++	+++	+++	+++	High
Brotman 2009 ⁹⁴	No	+++	+	+++	+	+++	+++	Medium
Brunson and Miller 2009 ¹⁵⁰	++	+++	+	+++	+++	+++	+++	High
Burnett 1998 ¹⁴⁹	No	No	No	+	+	+	++	Low
Cousins 1997 ¹⁶²	Unclear	Unclear	Unclear	++	+	+++	++	Low
Devine 1995 ¹⁵¹	No	No	No	++	+	+++	+	Low
Fletcher <i>et al.</i> 2009 ⁹¹⁻⁹³	++	++	++	+++	+++	+++	+++	High
Gordon and Turner 2001 ⁹⁵	No	+	+++	+++	+	+++	+++	Medium
Haselswerdt and Lenhardt 2003 ¹⁵⁷	Unclear	+	No	+++	+	+++	+++	Medium

Study	Were steps taken to increase rigour in sampling?	Were steps taken to increase rigour in data collected?	Were steps taken to increase rigour in analysis of data?	Were findings grounded in data?	Findings in terms of breadth	Findings in terms of depth	To what extent are the perspectives of young people privileged?	Overall reliability/trustworthiness
Hosie 2007 ¹⁵⁸	+	++	Unclear	+++	+++	+++	+++	Medium
Keddie 2009 ¹⁵⁹	+	+	No	+++	+	+++	+++	Medium
Lundblad <i>et al.</i> 2010 ¹⁶⁰	No	+	+	+++	+++	+++	+++	Medium
Marsiglia <i>et al.</i> 2002 ⁹⁶	+	+	No	+++	+	+	+++	Low
Pike and Colquhoun 2009 ¹⁶³	Unclear	Unclear	Unclear	+++	+++	+++	+++	Low
Plano Clark <i>et al.</i> 2002 ¹⁵⁵	+	+++	+	+++	+++	+++	+++	High
Waldron 2005 ¹⁶¹	+	+	No	+++	+++	+++	+++	Medium
Wills <i>et al.</i> 2005 ¹⁵⁶	++	+	+++	+++	+	+++	+++	High
Wilson-Simmons <i>et al.</i> 2006 ¹⁵²	+	+++	+++	+++	+++	+++	+++	High

+, few steps were taken to account for this criterion (limited); ++, several steps were taken to account for this criterion (fair); +++, a thorough attempt was made to account for this criterion (good).

Appendix 7 Protocol

Protocol for a systematic review of the effects of schools and school environment interventions on health: evidence mapping and syntheses

Chris Bonell,^{1*} Angela Harden,^{2†} Helene Wells,¹ Farah Jamal,^{2†} Adam Fletcher,^{1†} Mark Petticrew,^{1†} James Thomas,^{3†} Margaret Whitehead,^{4†} Rona Campbell,^{5†} Simon Murphy,^{6†} and Laurence Moore,^{6†}

*Corresponding author: chris.bonell@lshtm.ac.uk

†Contributed equally

¹Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London WC1H 9SH, UK. Full list of author information is available at the end of the article.

Bonell et al. *BMC Public Health* 2011, **11**:453

<http://www.biomedcentral.com/1471-2458/11/453>

Abstract

Background

Schools may have important effects on students' and staff's health. Rather than treating schools merely as sites for health education, 'school-environment' interventions treat schools as settings which influence health. Evidence concerning the effects of such interventions has not been recently synthesised.

Methods/design

Systematic review aiming to map and synthesise evidence on what theories and conceptual frameworks are most commonly used to inform school-environment interventions or explain school-level influences on health; what effects school-environment interventions have on health/health inequalities; how feasible and acceptable are school environment interventions; what effects other school-level factors have on health; and through what processes school-level influences affect health. We will examine interventions aiming to promote health by modifying schools' physical, social or cultural environment via actions focused on school policies and practices relating to education, pastoral care and other aspects of schools beyond merely providing health education. Participants are staff and students age 4–18 years. We will review published research unrestricted by language, year or source. Searching will involve electronic databases including Embase, ERIC, PubMed, PsycINFO and Social Science Citation Index using natural-language phrases plus reference/citation checking. Stage 1 will map studies descriptively by focus and methods. Stage 2 will involve additional inclusion criteria, quality assessment and data extraction undertaken by two reviewers in parallel. Evidence will be synthesised narratively and statistically where appropriate undertaking subgroup analyses and meta-regression and where no significant heterogeneity of effect sizes is found, pooling these to calculate a final effect size.

Discussion

We anticipate: finding a large number of studies missed by previous reviews; that non-intervention studies of school effects examine a greater breadth of determinants than are addressed by intervention studies; and that intervention effect estimates are greater than for school-based health curriculum interventions without school-environment components.

doi:10.1186/1471-2458-11-453

Cite this article as: Bonell et al.: Protocol for a systematic review of the effects of schools and school-environment interventions on health: evidence mapping and syntheses. *BMC Public Health* 2011 **11**:453.

Background

UK young people have among the worst health in Europe and there are marked inequalities in health across the social scale, with considerable implications for later health and economic costs [1,2]. Health education programmes are delivered through the school curriculum and aim to improve knowledge, develop skills and modify peer norms, and are now well-established in schools, addressing health behaviours such as smoking, drinking, drug use, sexual behaviour, physical activity and diet. However, numerous systematic reviews suggest such interventions have mixed and frequently disappointing results [3–9].

A complementary approach to curriculum-based health education is to change the school environment to promote health and wellbeing. The physical, social and cultural environment in which staff and students spend a high proportion of every weekday may have profound effects on their emotional and mental health, and opportunities to choose healthy lifestyles. Rather than treating schools merely as sites for health education, school-environment 'SE' interventions aim to modify how the school environment influences health. SE interventions can address health directly, for example: modifying school policies on smoking [10] etc.; improving catering [11]; or encouraging staff and students to walk or cycle to school [12]. Other actions aim to address factors such as disengagement and lack of social support that are risk factors for multiple adverse outcomes [13,14]. The latter include: increasing student participation in decision-making; providing staff with training on how to re-engage disaffected students; and encouraging students to take on new responsibilities such as becoming peer mediators [15]. These interventions take a 'socio-ecological' [16] approach to promoting health, whereby health is understood to be influenced not only by individual characteristics and behaviours, but also the wider social, cultural and economic context.

An important influence on the development of SE interventions has been the World Health Organisation's WHO framework for 'Health Promoting Schools' HPS [17]. This requires that schools simultaneously address their 'ethos' (i.e. school values and priorities, family/community involvement and curriculum). Some HPS have been rigorously evaluated but many have not [18]. Other trials have evaluated interventions which aim to modify the school environment to promote health but which are not explicitly informed by the HPS framework.

Evidence concerning the effects of SE interventions has not been comprehensively synthesised and several reviews that have examined these interventions are now quite old. A decade-old systematic review, focused only on HPS interventions, identifying only 12 studies, four of which were randomised trials. It concluded HPS interventions are promising, especially for promoting healthy eating, reducing bullying and improving mental and social wellbeing [18]. Other systematic reviews have focused on SE interventions that aim to reduce violence and drug use (not explicitly informed by the HPS framework) [19–22]. No evidence syntheses have been done on the effects of SE interventions in important areas such as sexual health, alcohol or smoking.

There has also been no synthesis of evidence on intervention process. Process evaluations examine the planning, delivery and receipt of SE interventions, and are useful for informing decisions about the wider implementation of interventions [23,24]. A further gap concerns synthesis of evidence on the health effects of the normal school environment, i.e. in the absence of intervention. This is important because to date SE intervention studies appear to have addressed only some aspects of the school environment and neglected others, such as school leadership and approaches to learning. Examining the impacts of such factors on health outcomes is now a growing field of public-health research [25] which merits synthesis. Although such studies provide less certain causal inference than experimental studies, those aiming to minimise confounding and other sources of bias could be used to identify promising areas for future intervention studies. A few reviews of such non-evaluation studies have been conducted but these either examine only certain outcomes or are unsystematic. Systematic reviews of school-level influences on drug use [22] and

smoking [26] have concluded there is, respectively, emerging and good evidence that factors such as teacher–student relationships and teaching styles may influence health. One non-systematic review of multi-level studies examined a range of health outcomes and, despite missing several important studies, suggested that strong leadership and high expectations appear to influence various health outcomes [27]. Finally, qualitative research has also been used to explore how staff and students perceive their school environment, and the processes they see as influencing health [28]. This evidence would also be useful in informing future SE interventions but remains unsynthesized.

We will work in close collaboration with colleagues in the Universities of Bristol and Cardiff undertaking a Cochrane review updating the decade-old review of interventions following the HPS framework; protocol available on request. While they focus on HPS interventions, we will examine the broader set of SE interventions and the other forms of evidence described above.

Our research questions are as follows:

- RQ1: What theories and conceptual frameworks are most commonly used to inform SE interventions or explain school-level influences on health? What testable hypotheses do these suggest?
- RQ2: What are the effects of SE (interventions aiming to promote health by modifying the school physical, social or cultural environment via actions focused on school policies and practices relating to education, pastoral care, sport, extra-curricular activities, catering, travel to and from school and other aspects of school life) evaluated using experimental and quasi-experimental designs compared with standard school practices on health (physical and emotional/mental health and wellbeing; intermediate health measures such as health behaviours, body mass index, teenage pregnancy; and health promotion outcomes such as health-related knowledge and attitudes) and health inequalities among school staff and students age 4–18 years? What are their direct and indirect costs?
- RQ3: How feasible and acceptable are SE interventions? How does context affect this?
- RQ4: What are the effects of other school-level factors on health and health inequalities among school staff and students age 4–18 years examined via multi-level and ecological (school) designs?
- RQ5: Through what processes might these school level influences occur?

Methods/design

The review will follow existing general criteria for the good conduct and reporting of systematic reviews (e.g. the Centre for Reviews and Dissemination guidelines; Quality of Reporting of Meta analyses guidelines). It will be carried out in two stages: (1) a descriptive map of available research evidence (which will involve exhaustive searching, application of inclusion and exclusion criteria, detailed coding), plus a preliminary synthesis of theories and conceptual frameworks used to inform SE interventions or explain school-level influences (on health on which we will consult with stakeholders to inform priorities for stage 2); and (2) a series of indepth syntheses in which the available research will be quality assessed, relevant findings extracted, and statistical and narrative/qualitative methods applied to synthesise findings

Stage 1: identifying and describing studies

In stage 1 we will include reports, without restrictions on language, date or source, that address each of our research questions.

We will exclude the following:

1. General topic – not about health/wellbeing or disease (including studies solely focused on outcomes concerned only with education).
2. Setting/population – not about the students or staff of schools (i.e. serving those age 4–18 years).
3. Type of report – not reporting primary research, a review of research or a theory

4. Specific focus

- 4a. (for intervention primary studies) – about an intervention that is neither mainly delivered on the school site nor concerned with travel to and from schools (extracurricular interventions will be included unless excluded based on any of the criteria below); neither about an intervention aiming to promote health/wellbeing or prevent disease nor reporting on the health/wellbeing outcomes of an intervention; about an intervention only involving: health education, information or counselling (regardless of who delivers this); school nursing, clinics or health checks; or health-related goods (medication, contraception, micronutrients etc.), but interventions concerning school catering, sport or active transport would be included; about an intervention targeted only to some students on the basis of health-related needs (but interventions targeted on the basis of educational or social but not health needs would be included).
- 4b. (for non-intervention primary studies) – not a study of the effects of the school environment/school-level factors on health/wellbeing.
- 4c. (for reviews and theoretical research) – not a review or theoretical paper with a focus on the school environment, interventions addressing this or school-level effects.

5. Study type

- 5a. (for intervention (primary studies)) – not an empirical outcome evaluation or process evaluation.
- 5b. (for non-intervention (primary studies)) – not empirically examining SE influences on health/wellbeing; if the study is a quantitative study it will be excluded if it is not reporting on school level variables (but multilevel analyses including school-level analyses would be included), only reporting on school-level measures of student social (e.g. SES) or demographic (e.g. ethnicity) characteristics or students' social networks (but studies examining student–staff relationships would be included), or only reporting on school-level measures of health education (regardless of who delivers this), school-based clinical health services or interventions targeted on the basis of health-related needs.
- 5c. (for reviews or theoretical research) – not a systematic review with a focus on school environment interventions, interventions to address this or school-level effects AND does not propose an abstracted, generalizable way in which features of schools are causally related to student/staff health.

The type of studies sought by this review are not likely to be reliably indexed in databases with controlled vocabularies. Therefore a very sensitive search will be undertaken using multiple natural language phrases (see *Appendix A* for PubMed search strategy). The first 'core' search strategy consists of four sets of terms relating to setting, population, intervention/influence and outcomes. A second search uses a broader set of 'non-core' terms covering these same areas. Some additional intervention terms will be added to the key terms as a third search. The intention is to sift the first set very carefully while the second and third set will be sifted more quickly. The following databases will be searched in July–August 2010, with no limits on language or date: Australian Educational Index; British Educational Index; CAB Health; The Campbell Library; CINAHL; Cochrane Controlled Trials Database; Embase; ERIC; Health Management Information Consortium; BSS; PubMed; PsycINFO; Social Policy and Practice (includes Child Data & Social Care Online); Social Science Citation Index (Web of Knowledge); Sociological Abstracts; and Dissertation Abstracts/Index to Theses. Econlit and PAIS were also investigated but trial searches produced no new material.

We will also undertake an intensive process of reference-checking of relevant papers, not only those references cited in the papers, but also looking for those papers which cite our target papers (using Citation Indexing in Web of Knowledge) and the Related Citations facility in Medline.

Search results will be downloaded into EPPI-Reviewer 4 software for screening. An inclusion criteria worksheet will be prepared, and each reference screened. Three reviewers will undertake these sifts, initially all three sifting the same studies and meeting to compare answers in initial batches of at least three sets of 50 studies to ensure consistency and more batches if required until the disparities are less than 5%, after which sifting will be done individually.

Studies will be descriptively coded based on title and abstract where possible and on full report where necessary. Included studies will be described by applying a standardized classification system for health promotion research [29] supplemented by new codes. For an initial sample, two reviewers will code independently, compare notes and reach consensus drawing on a third reviewer where necessary. Guidance for reviewers will be refined to remove any ambiguities that arise. Subsequent coding will be done by one reviewer. We will thus develop our evidence map.

Alongside this descriptive mapping, we will undertake a preliminary review of literature addressing RQ1. This synthesis will aim to develop hypotheses to be tested in our stage-2 synthesis regarding RQ2–5. Our review of theory will use thematic synthesis methods [30]. At this stage, we will engage with stakeholders via a workshop involving professionals and parent-governors, and a meeting involving young people. Each of these will review our evidence map and theory synthesis and provide comments that we will use these to inform our setting of hypotheses to be examined in stage 2. Additionally, if we identify a body of evidence of a size incommensurate with the planned scale of this evidence synthesis, we will also consult with these groups to determine priorities for stage 2.

Stage 2: In-depth syntheses addressing each research question

The final scope of the in-depth syntheses will be informed by our descriptive map, theory synthesis and stakeholder consultation. We will restrict in-depth syntheses to the best available evidence. Inclusion criteria relating to methodological quality will be applied to minimise bias. Where relevant these will be applied to each outcome and not merely to overall studies. Draft methodological inclusion criteria for stage 2 are as follows:

- RQ1: Not applicable: already synthesised in stage 1.
- RQ2: Prospective design with comparison groups; predetermined outcomes; control for clustering; control of confounding; no over-adjustment for potential mediators; and reporting on attrition, overall and by group (we will include in the review studies with >30% overall attrition, or >10% between group differences in attrition, but may exclude these from meta-analyses).
- RQ3. Process evaluations will not be excluded on the basis of quality but will be quality-assessed and their findings weighted (see below).
- RQ4: Control for clustering; control of school-compositional confounders; no over-adjustment for potential mediators; and reporting on attrition (again we may exclude studies with >30% attrition from meta-analyses). If sufficient studies, we will restrict our attention to multi-level, longitudinal studies which can better control for individual-level confounding and for reverse causality.
- RQ5: Qualitative studies will also not be excluded on the basis of quality but will be quality-assessed and their findings weighted (see below).

As in stage 1, criteria will be piloted prior to application. To help assure the review's quality at this stage, pairs of reviewers will first work independently and then compare their decisions before reaching consensus for all reports reviewed, involving a third reviewer where necessary. We will collect detailed data from, and describe, the included studies addressing RQs2–5. For all studies we will extract data on: study research questions/hypotheses; study site and population; sampling; data collection methods; analysis methods; results; and authors' conclusions. Additional data to be extracted for various study types are listed below.

- Quantitative studies addressing RQs 2 and 4: methods of adjustment for clustering; confounders and methods to control these; attrition rates overall and by study arm (RQ2 only); outcome measures; and effect size estimates overall and by population socio-economic, sex and ethnic sub-group and measures of confidence/significance.
- Economic studies addressing RQ2: (depending on what studies are found): intervention costs and indirect resource use; basis, assumptions or perspective taken regarding cost estimates; and (if available) economic measures of cost-effectiveness. In addition, we will extract other relevant data on study design and methods as per those listed above for quantitative studies.

- Qualitative studies addressing RQs 3 and 5: the rationale for the sampling method used; the range of stakeholder perspectives explored; and the transparency of reporting methods and data. For process evaluations we will also examine: part of process examined (planning, delivery, receipt); aspect of process examined (feasibility, fidelity/quality, coverage/accessibility, acceptability, appropriateness/fit with measured/perceived need); and aspect of intervention context examined (e.g. socio-demographic, policy, institutional capacity and collaboration, professional capacity). We have previously developed a tool for examining intervention context [31] which will be considered for use in this review, suitably adapted.

The quality of process evaluations and other qualitative research will be assessed according to a set of recently developed criteria used in an HTA-funded review of school-based interventions [32]. Reviewers will assess studies according to: the appropriateness of the sampling strategy to the evaluation aims; the rigour and, where appropriate, flexibility of data collection; the systematic and comprehensive nature of data analysis; whether findings are grounded in/supported by the data; whether the findings are of sufficient depth and breadth; and whether the perspectives of those involved in planning, delivering and receiving the interventions are adequately examined. A final step in the quality assessment of qualitative studies will be to assign studies two types of 'weight of evidence'. Firstly, reviewers will be asked to assign a weight (low, medium or high) to rate the reliability or trustworthiness of the findings (the extent to which the methods employed were rigorous/could minimise bias and error in the findings). Secondly, reviewers will also be asked to assign an additional weight (low, medium, high) to rate the usefulness of the findings for shedding light on factors relating to the research questions. Guidance will be given to reviewers to help them reach an assessment on each criterion and the final weight of evidence. Similarly, assessment and weighting of the methodological quality of any cost, economic evaluations and econometric studies that we find will be informed by application of existing methods and checklists [33,34].

In synthesising the evidence regarding RQs 2 and 4 we will undertake statistical meta-analysis when studies are sufficiently homogenous in terms of interventions (RQ2) and measures (RQs 2 and 4). Statistical heterogeneity of effects will be assessed using Chi-square tests and the magnitude of statistical heterogeneity will be assessed using the I^2 statistic. We will undertake subgroup analyses and meta-regression [35] and where no significant heterogeneity of effect sizes is found, these will be pooled to calculate a final effect size. While these analyses may enable us to hypothesise as to possible causes of differences between studies' findings, some heterogeneity is likely to remain, and any statistical analysis will be accompanied by a narrative synthesis.

Where data allow, our meta-analyses will aim to test hypotheses generated from our preliminary synthesis addressing RQ1. The use of a priori hypotheses from RQ1 will: give us an empirical justification for hypothesising that a given concept might impact on study findings; protect us from 'dredging' the data for spurious statistically significant results; and enable us to critique the selection of covariates that are employed in our included studies.

If the number of outcomes for which meta-analyses is possible exceeds the capacity of this project, we will focus on those outcomes prioritised by our stakeholder meeting. Meta-analysis and subgroup analysis will be conducted using EPPI-Reviewer with Stata 11 being used for any meta regression. As we anticipate that outcomes will be measured using a range of measurement tools, standardisation of results will be required in the form of standardised mean difference. We also anticipate that most of the studies addressing RQ2 will have used cluster randomised controlled trials, and most of those addressing RQ4 will have used multi-level or ecological (school) designs. We will draw on relevant methods [36] to calculate effects sizes from such studies. We will apply an 'equity lens' [37] to the effectiveness analysis (conducting sub-group analyses employing meta regression to examine any differences in impact according to socio-economic status, sex or ethnicity) in order to explore the potential impact of school-environment interventions on health inequalities. The precise hypotheses to be tested in these analyses will be determined by our theory synthesis.

Our synthesis of economic evaluations regarding RQ2 will be guided by what evidence we find. Measures of costs and (if available) indirect resource use and cost-effectiveness will be summarised using tables. If

measures of resource use are judged sufficiently homogeneous across studies, and applicable or transferable to the UK context, these will be synthesised using statistical meta-analysis [33]. Measures of costs, indirect resource use and cost-effectiveness collected from studies conducted outside the UK and/or in previous years will be adjusted for currency and inflation to the current UK. These data will be used to inform a narrative synthesis of the principal results of economic analyses, a commentary on economic aspects of school-environment interventions, and the applicability of collected economic evidence to the UK.

Findings from qualitative studies addressing RQs 3 and 5 will be synthesised using narrative methods [30,38,39]. Detailed evidence tables will be prepared to describe the methodological quality of each study, details of the intervention or aspect of schools examined, study site/population and findings. Two reviewers will read and re-read data contained within the evidence tables, apply codes and memos to capture the content of the data, and then group and organise codes into higher-order themes. These themes will be used to generate an explanatory framework to address RQs 3 and 5.

Published reports may be incomplete in a wide range of ways. For example: they may not report sufficient detail about their participants for our equity analysis; they may not present information on all the outcomes that were measured (possibly resulting in outcome reporting bias); they may not provide sufficient information about the intervention for accurate characterisation; and they may not report the necessary statistical information for the calculation of effect sizes. In all cases where there is a danger of missing data affecting our analysis, we will contact authors of papers wherever possible to request additional information. Where this process fails to provide the necessary detail (either because we cannot contact the authors, or they are unable to provide the information we need), we will need to use our judgement as to the most appropriate way forward. Statistical information, such as standard deviations and intra-cluster correlation co-efficients can be imputed from similar studies. We will use imputation where necessary – and defensible – and undertake sensitivity analyses to assess the impact of a range of possible values where this is done. In other instances of missing data (such as missing population information) it may not be possible to include a study in a particular analysis if, for example, it is impossible to classify the population using our equity tool.

Finally, we will draw on our five individual syntheses to produce a draft report. We will then organise stakeholder workshops with professionals and parent-governors, and young people to review our key findings and conclusions. Taking on board the views expressed by stakeholders, we will then finalise our technical report and executive summary, and begin disseminating the research via other means.

Discussion

We anticipate: finding a large number of studies missed by previous reviews; that non-intervention studies of school effects examine a greater breadth of determinants than are addressed by intervention studies; and that intervention effect estimates are greater than for school based health curriculum interventions without school environment components.

Additional material

Acknowledgements

This project is funded from a grant by the National Institute for Health Research Public Health Research Programme (grant 09/3002/08). We would like to thank Val Hamilton, Carol Vigurs, Beki Langford and Jeff Brunton for their contribution to the project. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR PHR programme or the Department of Health.

Author details

¹Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London WC1H 9SH, UK.

²Institute for Health and Human Development, University of East London, Water Lane, London E15 4LZ, UK.

³Social Science Research Unit, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL, UK.

⁴School of Population, Community and Behavioural Sciences, University of Liverpool, Whelan Building, Quadrangle, Brownlow Hill, Liverpool L69 3GB, UK.

⁵Department of Social Medicine, University of Bristol, 39 Whatley Road, Bristol BS8 2PS, UK.

⁶Cardiff Institute for Society, Health and Ethics, University of Cardiff, 1–3 Museum Place, Cardiff CF10 3BD, UK.

Authors' contributions

CB conceived and designed the study. AH contributed to the design of the study and in particular the methods relating to review and synthesis of qualitative data. HW and FJ contributed to the design of the sifting and data extraction processes. AF and MP contributed to the design of the synthesis of theory and qualitative data. JT contributed to the planned use of EPPI Reviewer software and to statistical methods. MW contributed to developing the background and planned outputs of the review. RC, SM and LM contributed to how this review inter-relates to a Cochrane review of multicomponent 'Health Promoting Schools' interventions. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 10 February 2011 Accepted: 9 June 2011

Published: 9 June 2011

References

1. Donaldson L: *Tackling the health of the teenage nation: Chief Medical Officer's Annual Report 2007*. London: Department of Health, England; 2008.
2. UNICEF: *An overview of child wellbeing in rich countries: a comprehensive assessment of the lives and wellbeing of children and adolescents in the economically advanced nations*. Florence: Innocenti Research Centre; 2007.
3. Harden A, Oakley A, Oliver S: Peer – delivered health promotion for young people: A systematic review of different study designs. *Health Education Journal* 2001;**60**:339–53.
4. Oliver S, Kavanagh J, Caird J, Lorenc T, Oliver K, Harden A, Thomas J, Greaves A, Oakley A: *Health promotion, inequalities and young people's health: a systematic review of research*. London: EPPI – Centre, Social Science Research Unit, Institute of Education, University of London; 2008.
5. DiCenso A, Guyatt G, Willan A, Griffith L: Interventions to reduce unintended pregnancies among adolescents: systematic review of randomised controlled trials. *Br Med J* 2002;**324**:1426–34.

6. Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P: School – based prevention for illicit drugs use. *Cochrane Database of Systematic Reviews* 2005;**2**:CD003020.
7. Foxcroft DR, Ireland D, Lowe G, Breen R: Primary prevention for alcohol misuse in young people. *Cochrane Database of Systematic Reviews*. *Cochrane Database of Systematic Reviews* 2002;**3**(2): CD003024, Art No CD003020.
8. Thomas R, Perera R: School – based programmes for preventing smoking (Cochrane Review). *Cochrane Database of Systematic Reviews* **3**:CD001293.
9. Wells J, Barlow J, Stewart-Brown S: A systematic review of universal approaches to mental health promotion in schools. *Health Education* 2003;**103**:197–220.
10. Cuijpers P, Jonkers R, de Weerd I, de Jong A: The effects of drug abuse prevention at school: the ‘Healthy School and Drugs’ project. *Addiction* 2002;**97**:67–73.
11. Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, Oakley A: *Young People and Healthy Eating: A systematic review of barriers and facilitators*. London: EPPI – Centre, Social Science Research Unit; 2001.
12. Rees R, Harden A, Shepherd J, Brunton G, Oliver S, Oakley A: *Young People and Physical Activity: A systematic review of barriers and facilitators*. London: EPPI – Centre, Social Science Research Unit; 2001.
13. Catalano RF, Haggerty KP, Oesterle S, Fleming CB, Hawkins JD: The importance of bonding to school for healthy development: findings from the Social Development Research Group. *J School Health* 2004;**74**(7):252–61.
14. Bond L, Carlin JB, Thomas L, Rubin K, Patton G: Does bullying cause emotional problems? A retrospective study of young teenagers. *Br Med J* 2001;**323**:480–84.
15. Bonell C, Sorhaindo A, Strange V, Wiggins M, Allen E, Fletcher A, Bond L, Flay B, Patton G, Rhodes T: A pilot whole – school intervention to improve school ethos and reduce substance use. *Health Education* 2010;**110**(4):252–272.
16. Dahlgren G, Whitehead M: *Policies and Strategies to Promote Social Equity in Health*. Stockholm: Institute for Futures Studies; 1991.
17. Young I, Williams T: *The Healthy School*. Edinburgh: Scottish Health Education Group; 1989.
18. Lister-Sharp DJ, Chapman S, Stewart-Brown S, Sowden A: Health promoting schools and health promotion in schools: two systematic reviews. *Health Technology Assessment* 1999;**3**(22).
19. Hahn R, Fuqua-Whitley D, Wethington H, Lowy J, Crosby A, Fullilove M: Effectiveness of universal school based programs to prevent violent and aggressive behaviour: a systematic review. *Am J Prev Med* 2007;**33**(2 Suppl):S114–S29.
20. Vreeman RC, Carroll AE: A systematic review of school – based interventions to prevent bullying. *Arch Pediatr Adolesc Med* 2007;**161**(1):78–88.
21. Wilson SJ, Lipsey MW: School – based interventions for aggressive and disruptive behaviour: update of a meta – analysis. *Am J Prev Med* 2007;**33**(2 Suppl):S130–S43.
22. Fletcher A, Bonell C, Hargreaves J: School effects on young people’s drug use: a systematic review of intervention and observational studies. *J Adolesc Health* 2008;**42**(3):209–20.
23. Oakley A, Strange V, Bonell C, Allen E, Stephenson J: Integrating process evaluation in the design of randomised controlled trials of complex interventions: the example of the RIPPLE Study. *Br Med J* 2006;**332**:413–6.
24. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M: Developing and evaluating complex interventions: the new Medical Research Council guidance. *Br Med J* 2008;**337**:a1655.

25. Bonell C, Fletcher A, McCambridge J: Improving school ethos may reduce substance misuse and teenage pregnancy. *Br Med J* 2007;**334**(7594):614–6.
26. Aveyard A, Markham WA, Cheng KK: A methodological and substantive review of the evidence that schools cause pupils to smoke. *Soc Sci Med* 2004;**58**:2252–65.
27. Sellström E, Bremberg S: Is there a 'school effect' on pupil outcomes? A review of multilevel studies. *J Epidemiol Community Health* 2006;**60**(2):149–55.
28. Fletcher A, Bonell C, Sorhaindo A, Strange V: How might schools influence young people's drug use? Development of theory from qualitative case – study research. *J Adolesc Health* 2009;**45**(2):126–32.
29. Peersman G, Oliver S, Oakley A: *EPPI – Centre Review Guidelines: Data Collections for the EPIC Database*. London: EPPI – Centre, Social Science Research Unit, Institute of Education, University of London; 1997.
30. Thomas J, Harden A: Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology* 2008;**8**:45.
31. Egan M, Bambra C, Petticrew M, Whitehead M: Reviewing evidence on complex social interventions: development and testing of a new tool for appraising implementation. *J Epidemiol Community Health* 2009;**63**:4–11.
32. Shepherd J, Kavanagh J, Picot J, Cooper K, Harden A, Barnett-Page E, Jones J, Clegg A, Hartwell D, Frampton G, Price A: The effectiveness and cost effectiveness of behavioural interventions for the prevention of sexually transmitted infections in young people aged 13 to 19: a systematic review and economic evaluation. *Health Technology Assessment* 2010;**14**(7):1–206.
33. Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M, Mallender J, McDaid D, Vale L, Walker D: Incorporating economics evidence. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.01*. Edited by: Higgins JPT, Green S. Oxford: The Cochrane Collaboration; 2008.
34. Rice N, Godfrey C, Slack R, Sowden A, Worthy G: *A systematic review of the effects of price on the smoking behaviour of young people: Final report*. York: Public Health Research Consortium; 2009.
35. Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, Wong G, Sheikh A: Can We Systematically Review Studies That Evaluate Complex Interventions? *PLoS Med* 2009;**6**(8):e1000086.
36. White I, Thomas J: Standardised mean difference in individually – randomised and cluster – randomised trials, with applications to meta – analysis. *Clinical Trials* 2005;**2**:141–51.
37. Kavanagh J, Oliver S, Lorenc T, Caird J, Tucker H, Greaves A, Thomas J, Oakley A: School – based cognitive behavioural interventions: a systematic review of effects and inequalities. *Health Sociology Review* 2009;**18**:61–78.
38. Arai L, Roen K, Roberts H, Popay J: It might work in Oklahoma but will it work in Oakhampton? Context and implementation in the effectiveness literature on domestic smoke detectors. *Inj Prev* 2005;**11**:148–51.
39. Noyes J, Popay J, Garner P: What can qualitative research contribute to a Cochrane systematic review of DOT for promoting adherence to tuberculosis treatment? Qualitative Research and Systematic Reviews workshop; 28 – 29 June; Continuing Professional Development Centre, University of Oxford 2005.

Pre-publication history

The pre-publication history for this paper can be accessed here: <http://www.biomedcentral.com/1471-2458/11/453/prepub>

Appendix A Searches

TABLE 1 Search (1): Medline core search strategy (29 July 2010)

Set	Searches	Results
1.	health promoting school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(137)
2.	(healthy school or healthy schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(448)
3.	comprehensive school* health program*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(48)
4.	(coordinated school* health program* or co-ordinated school* health program*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(55)
5.	1 or 2 or 3 or 4	(659)
6.	(whole school or school wide or schoolwide).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(207)
7.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(165219)
8.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2479458)
9.	7 and 8	(100116)
10.	ethos.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(640)
11.	(school* adj5 climate).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(215)
12.	exp Climate Change/ or exp Climate/	(86577)
13.	11 not 12	(195)
14.	(school* adj5 environment).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1335)
15.	(school* adj5 culture).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(207)
16.	(school* adj3 manag*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(556)
17.	(school* adj3 leader*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(226)
18.	(school* adj5 or gani?ation).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(341)

continued

TABLE 1 Search (1): Medline core search strategy (29 July 2010) (*continued*)

Set	Searches	Results
19.	(school* adj5 aggregat*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(59)
20.	(school* adj5 governance).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(27)
21.	education*context*.mp.	(143)
22.	(interschool variation* or inter-school variation*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4)
23.	(interschool differen* or inter-school differen*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1)
24.	(interschool inequalit* or inter-school inequalit*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(0)
25.	(school difference* or differen*between school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(77)
26.	(school* adj2 level).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1253)
27.	(school* adj3 varia*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(610)
28.	school influence*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(40)
29.	((school* adj3 effect (or school) adj3 effects)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(431)
30.	rest or *justice.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(39)
31.	Schools/og [Organization & Administration]	(619)
32.	Health Promotion/	(39229)
33.	31 and 32	(89)
34.	10 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 33	(5808)
35.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1430864)
36.	(wellbeing or wellbeing).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29403)
37.	infection*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1087334)
38.	disease*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3636128)
39.	emotion*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(91977)

TABLE 1 Search (1): Medline core search strategy (29 July 2010) (*continued*)

Set	Searches	Results
40.	mental.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(310841)
41.	psychiatr*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(214115)
42.	anxi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(116132)
43.	depress*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(307948)
44.	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43	(5846882)
45.	Juvenile Delinquency/	(6383)
46.	Violence/pc [Prevention & Control]	(3758)
47.	(violen* or delinquen* or aggress* or bully* or bullies or bullied).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(145752)
48.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(650276)
49.	victimi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2928)
50.	45 or 46 or 47 or 48 or 49	(785497)
51.	Alcohol Drinking/	(41336)
52.	Smoking/ or Marijuana Smoking/	(100310)
53.	Substance-Related Dis or ders/pc [Prevention & Control]	(6277)
54.	(smok* or tobacco).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(201065)
55.	cigarett*.mp.	(39495)
56.	(illegal drug* or illicit drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(10677)
57.	(drug* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(51659)
58.	(drug* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(22736)
59.	57 not 58	(28923)
60.	(drug* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29097)
61.	(substance* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(12767)
		continued

TABLE 1 Search (1): Medline core search strategy (29 July 2010) (*continued*)

Set	Searches	Results
62.	(substance* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1946)
63.	61 not 62	(10821)
64.	(substance* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(33382)
65.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(14873)
66.	(alcohol or binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(168866)
67.	51 or 52 or 53 or 54 or 55 or 56 or 59 or 60 or 63 or 64 or 65 or 66	(416873)
68.	(healthy eating or nutrition or diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(334194)
69.	(obesity or overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(135458)
70.	(bodyweight or body weight or bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(309228)
71.	(physical exercise or physical*activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(45281)
72.	68 or 69 or 70 or 71	(681558)
73.	active commuting to school.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(34)
74.	(active transport not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2668)
75.	(walking bus or walking buses or school travel plan*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1)
76.	((walk* adj2 school*) or cycl*) adj2 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(145)
77.	73 or 74 or 75 or 76	(2823)
78.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	(20442)
79.	exp Contraceptive Devices/ut [Utilization]	(3777)
80.	Pregnancy in Adolescence/	(5869)
81.	SexualBehavior/	(34372)
82.	Sexually Transmitted Diseases/ep, pc [Epidemiology, Prevention & Control]	(9928)
83.	(pregnan* or sexual).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(795766)

TABLE 1 Search (1): Medline core search strategy (29 July 2010) (*continued*)

Set	Searches	Results
84.	(HIV or chlamydia).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(227757)
85.	(condom* or contracepti*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(71857)
86.	78 or 79 or 80 or 81 or 82 or 83 or 84 or 85	(1038088)
87.	44 or 50 or 67 or 72 or 77 or 86	(7414552)
88.	9 and 34 and 87	(2943)
89.	6 and 87	(173)
90.	5 or 88 or 89	(3633)
91.	schools, dental/ or schools, medical/ or schools, nursing/ or schools, pharmacy/ or schools, public health/ or schools, veterinary/	(29874)
92.	medical school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(19144)
93.	91 or 92	(43387)
94.	90 not 93	(3518)

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010)

Set	Searches	Results
1.	health promoting school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(139)
2.	(healthy school or healthy schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(450)
3.	comprehensive school*health program*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(48)
4.	(coordinated school*health program* or co-ordinated school*health program*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(56)
5.	1 or 2 or 3 or 4	(664)
6.	(whole school or school wide or schoolwide).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(208)
7.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(165694)

continued

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
8.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2485450)
9.	7 and 8	(100468)
10.	ethos.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(641)
11.	(school* adj5 climate).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(218)
12.	exp Climate Change/ or exp Climate	(86839)
13.	11 <i>not</i> 12	(198)
14.	(school* adj5 environment).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1339)
15.	(school* adj5 culture).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(207)
16.	(school* adj3 manag*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(558)
17.	(school* adj3 leader*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(228)
18.	(school* adj5 organi?ation).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(341)
19.	(school* adj5 aggregat*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(60)
20.	(school* adj5 governance).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(27)
21.	education*context*.mp.	(145)
22.	(interschool variation* or inter-school variation*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4)
23.	(interschool differen* or inter-school differen*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1)
24.	(interschool inequalit* or inter-school inequalit*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(0)
25.	(school difference* or differen* between school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(78)
26.	(school* adj2 level).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1262)

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
27.	(school* adj3 varia*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(611)
28.	school influence*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(40)
29.	((school* adj3 effect) or school) adj3 effects).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(438)
30.	restor *justice.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(40)
31.	Schools/og [Organization & Administration]	(621)
32.	Health Promotion/	(39392)
33.	31 and 32	(89)
34.	10 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 33	(5841)
35.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1435980)
36.	(wellbeing or well being).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29542)
37.	infection*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1090528)
38.	disease*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3646240)
39.	emotion*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(92408)
40.	mental.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(311800)
41.	psychiatr*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(214674)
42.	anxi*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(116609)
43.	depress*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(308909)
44.	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43	(5863435)
45.	Juvenile Delinquency/	(6409)
46.	Violence/pc [Prevention & Control]	(3768)
47.	(violen* or delinquen* or aggress* or bully* or bullies or bullied).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(146311)

continued

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
48.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(652077)
49.	victim*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2958)
50.	45 or 46 or 47 or 48 or 49	(787823)
51.	Alcohol Drinking/	(41502)
52.	Smoking/ or Marijuana Smoking/	(100634)
53.	Substance-Related Disorders/pc [Prevention & Control]	(6299)
54.	smok* or tobacco.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(201762)
55.	cigarett*.mp	(39637)
56.	(illegal drug* or illicit drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(10721)
57.	(drug* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(51908)
58.	(drug* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(22831)
59.	57 not 58	(29077)
60.	(drug* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29189)
61.	(substance* adj2 use?).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(12891)
62.	(substance* adj2 (used or uses)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1949)
63.	61 not 62	(10942)
64.	(substance* adj2 (abus* or misus* or users)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(33519)
65.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(14930)
66.	(alcohol or binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(169462)
67.	51 or 52 or 53 or 54 or 55 or 56 or 59 or 60 or 63 or 64 or 65 or 66	(418331)
68.	(healthy eating or nutrition or diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(335062)
69.	(obesity or overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(136129)

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
70.	(bodyweight or body weight or bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(310258)
71.	(physical exercise or physical*activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(45514)
72.	68 or 69 or 70 or 71	(683716)
73.	active commuting to school.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(34)
74.	(active transport not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2671)
75.	(walking bus or walking buses or school travel plan*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1)
76.	((walk* adj2 school*) or cycl*) adj2 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(145)
77.	73 or 74 or 75 or 76	(2826)
78.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	(20453)
79.	exp Contraceptive Devices/ut [Utilization]	(3800)
80.	Pregnancy in Adolescence/	(5879)
81.	Sexual Behavior/	(34470)
82.	Sexually Transmitted Diseases/ep,pc [Epidemiology, Prevention & Control]	(9976)
83.	(pregnan* or sexual).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(797455)
84.	(HIV or chlamydia).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(228565)
85.	(condom* or contracepti*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(72004)
86.	78 or 79 or 80 or 81 or 82 or 83 or 84 or 85	(1040475)
87.	44 or 50 or 67 or 72 or 77 or 86	(7434634)
88.	9 and 34 and 87	(2966)
89.	6 and 87	(174)
90.	5 or 88 or 89	3660
91.	schools, dental/ or schools, medical/ or schools, nursing/ or schools, pharmacy/ or schools, public health/ or schools,veterinary/	(29917)
		continued

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
92.	medical school*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(19187)
93.	91 or 92	(43464)
94.	90 not 93	(3545)
95.	(multiintervention or multi-intervention).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(30)
96.	(noncurricul* or non-curric*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(15)
97.	(socioecolog* or socio-ecolog*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(264)
98.	(socioenvironment* or socio-environment*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(854)
99.	classroom management.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(44)
100.	(value added not (tax or VAT)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(794)
101.	past or al.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3566)
102.	((school* adj3 achievement*) or (school* adj3 attainment*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1059)
103.	((school* adj3 exam*) or (school adj3 (test or tests or testing or tested))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2883)
104.	(ecological level or ecological study or ecological studies or ecological design* or ecological measure* or ecological variable* or ecological analysis or school ecology).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]2368	(2368)
105.	((student*engagement or engaging students or pupil*engagement or engaging pupils or engaging school or (school* adj2 engagement) or (education* adj2 engagement)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(235)
106.	95 or 96 or 97 or 98 or 99 or 100 or 101 or 104 or 105	(8147)
107.	((school* adj3 qualif*) or (school* adj3 quality) or (school* adj3 inspect*) or (school* adj3 influence*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1125)
108.	((school* adj5 (policy or policies) or (school* adj3 rules) or (school* adj3 practices).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1648)

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
109.	((school* adj5 context* or (school* adj3 opportunit*) or (school* adj5 collective)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(787)
110.	((school* adj3 communit*) or (school* adj3 sttructur*) or (school* adj3 relation*) or (school* adj3 communicat*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3437)
111.	((school* adj3 structur*) or (school* adj3 security) or (school* adj3 safe*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(590)
112.	((school* adj3 expectation*) or (school* adj2 exclusion*) or (school inclusion or (between adj2 schools)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(661)
113.	102 or 103 or 107 or 108 or 109 or 110 or 111 or 112	(11145)
114.	((education* adj3 achievement*) or (education* adj3 attainment*) or (education* adj3 exam*) or (education* adj3 (tests or test or testing))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(6087)
115.	((education* adj3 qualif*) or (education* adj3 quality) or (education* adj3 (policy or policies))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4058)
116.	((education* adj3 opportunit*) or education* adj3 practices or (education* adj3 culture) or (education* adj3 manag*) or (education* adj3 leader*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(5904)
117.	((education* adj3 practices) or (education* adj3 communicat*) or (education* adj3 safe*) or (education* adj3 expectation*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3730)
118.	114 or 115 or 116 or 117	(19103)
119.	((teaching adj3 practices) or (teaching adj3 standard*) or (teaching adj3 style*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(895)
120.	((teaching adj3 method*) or (teaching adj3 differen*) or (teaching adj3 varia*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4857)
121.	119 or 120	(5643)
122.	(pupil-led or student led).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(53)
123.	aggregate* adj2 (data or reports or information).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1591)
		<i>continued</i>

TABLE 2 Search (2): Medline non-core search strategy (12 August 2010) (*continued*)

Set	Searches	Results
124.	(school* <i>size</i> or school restructur* or comprehensive school reform).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(72)
125.	122 or 123 or 124	(1716)
126.	106 or 113 or 118 or 121 or 125	(44608)
127.	9 and 87 and 126	(8235)
128.	127 not 93	(7767)
129.	128 not 94	(6793)

TABLE 3 Search (3): Medline additional search strategy (23 September 2010)

Set	Searches	Results
1.	(child* or adolescen* or youth or young people or teen* or student or pupil* or teacher* or teaching staff or school personnel or school staff or parent*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(2568402)
2.	(school or schools).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]167005	(167005)
3.	1 and 2	(113129)
4.	(active transport adj3 school*not cell*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(13)
5.	active commuting to school.mp. [mp=title, title, abstract, name of substance word, subject heading word, unique identifier]	(34)
6.	(walking bus or walking buses or walking school bus*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(7)
7.	school travel plan*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1)
8.	(walk* adj3 school*).ti.ab.	(130)
9.	((cycle adj3 school*) or cycling adj3 school*) or (bicycl* adj3 school*).ti.ab.	(85)
10.	4 or 5 or 6 or 7 or 8 or 9	(228)
11.	3 and 10	(213)
12.	((school* adj3 dinner*) or (school* adj3 lunch*) or (school* adj3 breakfast*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(646)

TABLE 3 Search (3): Medline additional search strategy (23 September 2010) (*continued*)

Set	Searches	Results
13.	breakfast club*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(8)
14.	(soft drink* machine* or snack* machine*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(6)
15.	12 or 13 or 14	(655)
16.	3 and 15	(595)
17.	school meals.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(234)
18.	(school* adj3 meal*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(354)
19.	18 not 17	(120)
20.	3 and 19	(111)
21.	16 or 20	(671)
22.	physical train*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(3762)
23.	"Physical Education and Training"/	(10602)
24.	(sport* adj3 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(512)
25.	(games adj3 school*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(28)
26.	22 or 23 or 24 or 25	(13770)
27.	3 and 26	(1462)
28.	health.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1448176)
29.	(wellbeing or well-being).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29911)
30.	(depress* or emotion* or mental or psychiatr* or anxi*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(811688)
31.	(infection* or disease*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4298493)
32.	28 or 29 or 30 or 31	(5905334)
33.	Violence/pc [Prevention & Control]	(3799)
34.	(bully* or bullied or bullies).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1193)
35.	(injur* or accident*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(656888)

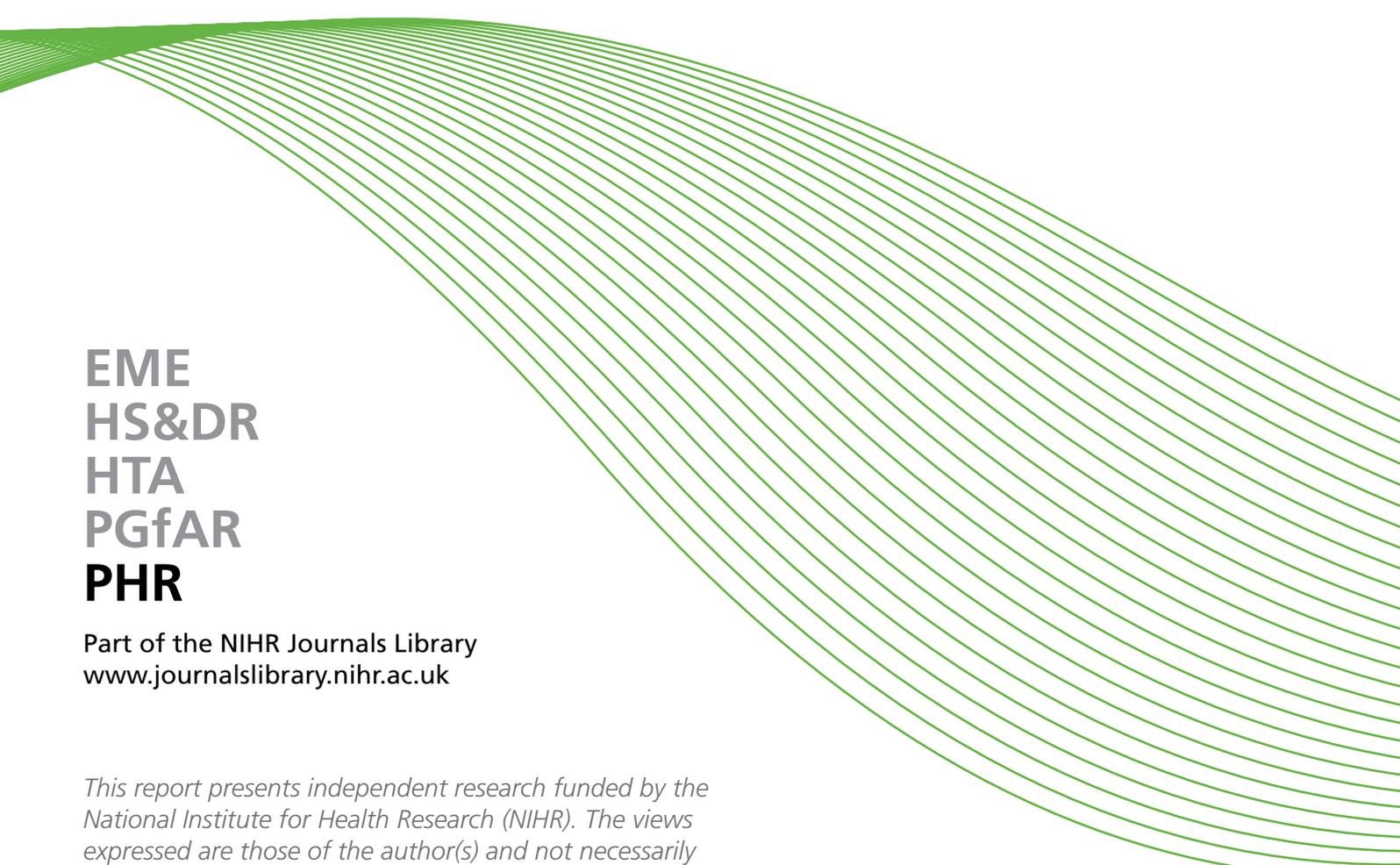
continued

TABLE 3 Search (3): Medline additional search strategy (23 September 2010) (*continued*)

Set	Searches	Results
36.	Alcohol Drinking/	(41859)
37.	Substance-Related Disorders/pc [Prevention & Control]	(6337)
38.	Smoking/ or Marijuana Smoking/	(101457)
39.	(substance* adj2 (user* or abuse* or misuse*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(33547)
40.	(smok* or tobacco).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(203627)
41.	cigarett*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(39976)
42.	(drug* adj2 (user* or abuse* or misuse*)).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(25626)
43.	(illicit drug* or illegal drug* or street drug*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(7310)
44.	(cannabis or marijuana).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(15055)
45.	(alcohol).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(167916)
46.	(binge).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(4669)
47.	(healthy eating or healthy-eating).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(1388)
48.	(obese or obesity).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(142720)
49.	(nutrition).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(118291)
50.	(diet).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(241314)
51.	(overweight or over-weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	23296
52.	(bodyweight or body weight).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	229595
53.	(bodymass or body mass).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	101322
54.	(physical exercise or physical*activ* or physical training*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(46166)
55.	47 or 48 or 49 or 50 or 51 or 52 or 53 or 54	(696215)
56.	Condoms/ut [Utilization]	(3065)

TABLE 3 Search (3): Medline additional search strategy (23 September 2010) (*continued*)

Set	Searches	Results
57.	HIV Infections/ep, pc [Epidemiology, Prevention & Control]	(36132)
58.	teenage pregnancy.mp. or Pregnancy in Adolescence/	(6203)
59.	Sexual Behavior/	(34770)
60.	Sexually Transmitted Diseases/ep, pc [Epidemiology, Prevention & Control]	(10056)
61.	sexual.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(137750)
62.	pregnan*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(678875)
63.	Acquired Immunodeficiency Syndrome/ep, pc [Epidemiology, Prevention & Control]	(20524)
64.	HIV*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(211093)
65.	chlamydia*.mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(20869)
66.	(condom or condoms or contracept*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(72313)
67.	56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 or 64 or 65 or 66	(1046418)
68.	(violen* or aggress* or victimi*).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(142137)
69.	33 or 34 or 35 or 68	(787707)
70.	((drug* adj2 use?) not (drug* adj2 (used or uses))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(29411)
71.	((substance* adj2 use?) (not substance* adj2 (used or uses))).mp. [mp=title, original title, abstract, name of substance word, subject heading word, unique identifier]	(11154)
72.	36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 70 or 71	(419192)
73.	32 or 55 or 67 or 69 or 72	(7482875)
74.	27 and 73	(1144)
75.	11 and 73	(183)
76.	21 and 73	(576)

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library