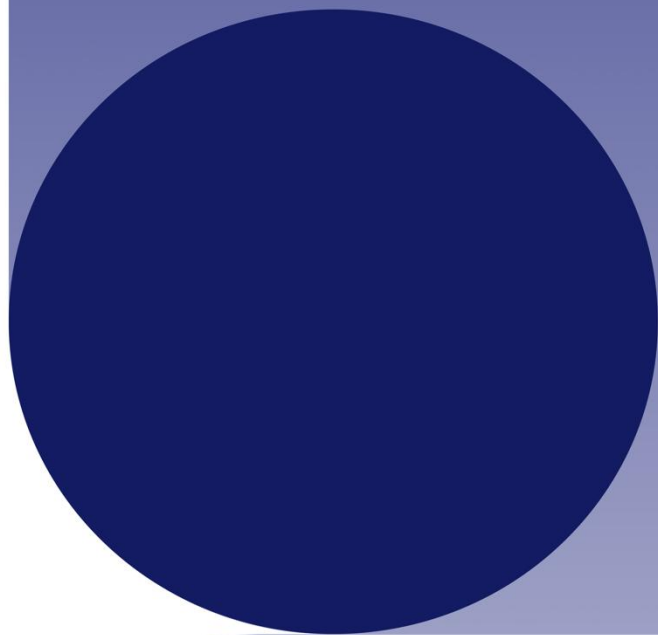




Education
Endowment
Foundation



Understanding school and classroom climate and its impact on pupil outcomes

Protocol for Review 2: Understanding and Conceptualising School and Classroom Climate

Darren Moore, Eleni Dimitrellou, Yuanya Zhang, Josie
Scammell, Simon Benham-Clarke, Katherine Gulliver, Alison
Bethel, Tom Ralph and Neil Harrison



University
of Exeter



The Education Endowment Foundation is an independent charity dedicated to breaking the link between family income and education achievement. We support schools, colleges, and early years settings to improve teaching and learning for 2 – 19-year-olds through better use of evidence.

We do this by:

- **Summarising evidence.** Reviewing the best available evidence on teaching and learning and presenting in an accessible way.
- **Finding new evidence.** Funding independent evaluations of programmes and approaches that aim to raise the attainment of children and young people from socio-economically disadvantaged backgrounds. Putting evidence to use.
- **Putting evidence to use.** Supporting education practitioners, as well as policymakers and other organisations, to use evidence in ways that improve teaching and learning.

We were set up in 2011 by the Sutton Trust partnership with Impetus with a founding £125m grant from the Department for Education. In 2022, we were reendowed with an additional £137m from government, allowing us to continue our work until at least 2032.

For more information about the EEF or this report please contact:



Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
SW1P 4QP



0207 802 1653



info@eefoundation.org.uk



www.educationendowmentfoundation.org.uk



Department
for Education



Table of contents

Background and review rationale.....	3
Review 1 interim findings.....	4
School climate.....	4
Classroom climate.....	4
Review 2 objectives.....	5
Methodology.....	6
Inclusion and exclusion criteria for the review.....	7
Search strategy for identification of studies	9
Selection of studies.....	10
Data extraction and management	10
Appraisal of included studies.....	11
Data synthesis.....	11
Reporting	19
Registration.....	19
Timeline.....	19
Team	19
Conflicts of interest	20
References	21
Appendices.....	23
Appendix 1: Review 1 School climate conceptualisation	23
Appendix 2: Review 1 School climate logic model.....	23
Appendix 3: Review 1 classroom climate conceptualisation.....	24
Appendix 4: Review 1 Classroom climate logic model.....	24
Appendix 5: Example Search Strategy (ERIC).....	25
Appendix 6: Snippet of example matrix of pooled effect sizes and confidence intervals	27
Appendix 7: Detailed Review 2 Timeline	28

Background and review rationale

This programme of work comprises two complementary systematic reviews examining school climate (SC) and classroom climate (CC), aiming to advance both conceptual clarity and evidence-informed practice. Review 1 adopted a theory-based approach to synthesise existing definitions, domains, and dimensions of SC and CC, culminating in a conceptual framework and logic models that articulate hypothesised causal pathways to outcomes. Building on this foundation, Review 2 will undertake a systematic review of interventions designed to improve SC and/or CC (SC/CC), with a focus on evaluating their impact on pupil outcomes and clarifying the intervention components, contextual conditions, and implementation factors that shape effectiveness. Together, the reviews are intended to generate insights that are both theoretically robust and directly relevant to policy and practice in school settings. This document provides the protocol which guides the work for Review 2.

Despite a substantial and longstanding body of research on SC in particular, the evidence base remains limited in its practical utility (Grazia & Molinari, 2021). Conceptualisation and measurement are highly variable, and intervention studies differ widely in scope, focus, and design, making synthesis and comparison challenging (Baumsteiger et al., 2023). While positive associations between climate and pupil outcomes are well established, there is comparatively limited robust evidence on causal impacts, including how climate-focused interventions produce change, for whom they are most effective, and under what conditions. Critical gaps remain in understanding mechanisms of change, differential effects across pupil groups, developmental appropriateness, and the implementation factors that affect feasibility, sustainability, and scalability in schools. As a result, schools and policymakers lack clear, actionable guidance on how to recognise, select and implement effective climate interventions, underscoring the need for a focused and systematic synthesis of impact evidence.

Review 1 adopted a theory-based systematic review design, drawing on guidance from White (2018), to examine how SC and CC are conceptualised across the literature. Existing review literature was used given the size of the evidence base and preference for established and synthesised conceptualisations of SC and CC. To ensure that original conceptualisations were not overlooked, reviews were supplemented with primary research studies and discussion papers that proposed new or adapted theories, models, frameworks or measures of SC and CC. The review 1 synthesis included three main components which were conducted for SC and CC, given little literature that compares and contrasts both constructs. The initial synthesis involved a systematic review of definitions to identify, collate, and critically appraise how SC and CC are defined across the literature. Drawing on Gillett-Swan et al. (2025), definitions were analysed beyond simple aggregation to identify shared features, points of divergence, and gaps. Each definition was also assessed against adapted criteria based on Joseph's rules for sound definition, allowing the review to evaluate clarity, precision, and explanatory power. This structured comparison supported the development of an integrative definition for each of SC and CC. Relatedly we also synthesised the domains of SC and CC featured in definitions, theories, models and frameworks, as well as measures. Therefore, our conceptualisation of SC and CC defined the terms and indicated what each construct is made up of through commonalities in the theoretical literature.

Causal chain analysis (CCA) was used to support the construction of initial logic models that explain how each of SC and CC are theorised to influence pupil outcomes. Following guidance from White (2018) and Kneale et al. (2018), this process began with the definitions produced in the prior synthesis and identified distal outcomes (e.g., academic achievement, wellbeing), intermediate outcomes (e.g., engagement and behaviour), and intervention and school setting contextual factors. Literature used included both theory hypothetically linking SC and CC with outcomes, and theoretical models tested in primary research, for instance assessing whether SC mediates the link between participation in school decision making and academic engagement and motivation. Evidence from included studies was used to map hypothesised causal pathways, articulate underlying assumptions, and identify relevant contextual moderators. The working logic models indicate gaps and uncertainties in the evidence base drawn from reviews and more theoretical papers (e.g. lack of contextual factors and teacher outcomes related to CC). The logic models provide the foundation which will be tested and refined in Review 2 using the evidence from intervention evaluations to identify causal pathways and explore which gaps persist in the intervention literature.

Review 1 interim findings

Review 1 located and screened 10,550 records and led to 454 included studies. These included studies were predominantly focused on school climate (n=209), with only 53 studies focused on classroom climate and only six studies directly addressing both in the same paper. Many studies therefore focused on overlapping terms including school culture, school organisational climate, classroom environment, school environment, school connection.

Because separate studies nearly always focused on SC versus CC, we synthesised separate definitions, domains and logic models for each construct. Of the 209 studies focused on SC, 84 were reviews, 76 original theories, models or frameworks, and 83 original measures of SC. 94 studies contributed to the SC logic model synthesis, most often reviews (n=39) and analytical models (n=39). Of the 53 studies focused on CC, 8 were reviews, 13 original theories, models or frameworks, and 35 original measures of CC. Only 16 studies contributed to the CC logic model synthesis, a mix of each study type.

School climate

When synthesising conceptualisations of SC we drew on 129 different definitions. Our synthesised definition of school climate is as follows:

School climate is the quality and character of school life, reflecting the shared perceptions and patterns of experiences among individuals in the school community, including pupils, all school staff and families. It is shaped by school culture – the deeper assumptions, values, beliefs and norms of the school – and school leadership. It is a multi-dimensional construct comprising relationships, teaching and learning, safety, and the school environment.

152 studies listed domains of SC, most often naming 3-6 domains. Four domains were seen in over half of these studies:

- Relationship/Community – 139 mentions
- Safety – 86 mentions
- Teaching and learning – 109 mentions
- Institutional environment – 90 mentions

See Appendix 1 for a diagram combining the definition and domains, indicating key sub domains from the literature.

The logic model for SC synthesises theoretical and review-level evidence of how SC is expected to influence outcomes. The model brings together three key elements: (1) interventions (including school-wide, behavioural, leadership, and related programmes), (2) factors influencing school climate (such as leadership, relationships, teaching practices, resources, and wider policy context), and (3) the core domains of school climate itself (relationships and community, safety and behaviour, teaching and learning, and the school environment). The logic model identifies proximal and distal outcomes, including engagement, behaviour, mental health, academic attainment, and teacher outcomes. See Appendix 2 for a draft of the SC logic model.

We also synthesised 78 studies focused on school culture to distinguish school climate as focused on perceptions and experiences, observable and dynamic, compared to school culture as focused on values and beliefs, underlying behaviour and more stable. School culture shapes school climate by providing the deeper assumptions, values, beliefs and norms through which school life is organised, interpreted and experienced.

Classroom climate

When synthesising conceptualisations of CC we drew on 31 different definitions. Our synthesised definition of classroom climate is as follows:

Classroom climate is a multidimensional construct representing the characteristics of classroom life. It reflects the shared perceptions and experiences of pupils, teachers and other school staff within a classroom and is shaped through their interactions and the physical environment. It comprises social-

emotional and relational dynamics, teaching and learning experiences, and behavioural and classroom organisation.

49 sets of domains of CC were seen across studies. Three domains were seen in over half of these:

- Relational and social-emotional – 40 mentions
- Teaching and learning – 30 mentions
- Behaviour and organisation – 25 mentions

See Appendix 3 for a diagram combining the definition and domains of CC, indicating key sub domains from the literature.

The logic model for CC synthesises theoretical and review-level evidence of how CC is expected to influence outcomes. The model brings together two key elements: (1) factors influencing school climate (such as teacher behaviour and group size), and (2) the core domains of CC itself (relational and social-emotional; teaching and learning; behavioural and organisational). The logic model identifies proximal and distal outcomes, including engagement, motivation, wellbeing and academic attainment. See Appendix 4 for a draft of the CC logic model.

Review 1 has established what SC and CC are and how they are conceptualised within the literature, providing a clear foundation for Review 2. It also indicates that there is far less literature on CC. However, Review 1 does not address whether and how climate can be meaningfully improved in practice. The next step is therefore to examine the extent to which interventions can improve SC and CC in real-world school settings, and what this means for pupil and teacher outcomes. Review 2 will focus on synthesising evidence from intervention studies to understand what works, for whom, and under what conditions, with particular attention to the roles of intervention design, context, and implementation in shaping effectiveness.

Review 2 objectives

This section outlines the review aims, objectives and research questions. Review 2 addresses the second of the two overarching aims of the research project:

We aim to advance understanding of SC and CC by

- (1) developing a clear, research-informed conceptual framework to address definitional ambiguity, and
- (2) identifying effective, evidence-based approaches that improve pupil and teacher outcomes, while exploring any differential effects by subgroups.

The overarching goal is to support practice and policy by generating practical, actionable recommendations for educators in England, grounded in robust evidence. This work recognises that SC and CC may both relate to aspects of school leadership and teaching quality, meaning findings should speak to different school staff.

Review questions that relate to the second systematic review with explanation are as follows:

RQ1: How are interventions targeting school and classroom climate designed and delivered? To what extent do they align with the conceptualisation of school and classroom climate established in Review 1?

This research question focuses on establishing what SC and CC interventions evaluated in the literature are like, how they are delivered and the extent to which the interventions fit the conceptualisation of SC/CC established in Review 1. This implies that we are not only including interventions that fit our conceptualisation derived from the literature.

RQ2: How effective are different types of school and classroom climate interventions in improving pupil outcomes, including attainment and non-cognitive outcomes (e.g., wellbeing, engagement, behaviour), as well as school staff outcomes?

Previously, this was an additional research question. However, it seems fundamental to answer how effective different types of interventions are (e.g. directly versus indirectly targeting climate) for the range of outcomes evaluated. The

wording has been changed to consider how effective different types of interventions are, rather than only focus on which individual interventions are effective.

RQ3: What conditions and strategies support the successful implementation and long-term sustainability of school and classroom climate interventions in diverse educational settings?

We have also upgraded this implementation research question to a main research question rather than an additional area of interest. Some conditions and strategies are likely to be considered as part of the next research question, but not all implementation evidence is likely to be causal (e.g. process evaluations).

RQ4: What school climate and classroom climate intervention components and contextual factors (and combinations of these components and factors) are causally linked to improvements in key outcomes, like behaviour, mental health and attainment?

This research question has not changed from previous plans. Causal chain analysis will be used to identify links between factors, intervention components and outcomes. Qualitative comparative analysis will be used to explore combinations of components and factors which predict effectiveness in key outcomes.

RQ5: How do the effectiveness and implementation of school and classroom climate interventions vary across different school phases, settings, and populations – particularly in relation to age, socio-economic disadvantage, SEND, Children in Care and other forms of inequality?

This equity analysis remains as planned. While we will be very interested in whether any subgroups might moderate effectiveness, we do not expect to be able to say for instance that primary or secondary SC/CC interventions are more effective given the diversity in interventions we are likely to see. But we will narratively describe any relevant findings across included studies.

RQ6: How does the research literature from across Review 1 and Review 2 help to answer key questions, such as:

- a) **What distinguishes school and classroom climate from each other and the following, often overlapping, terms: school connection, belonging, school ethos, authoritative school climate?**
- b) **Do interventions that directly target school climate or classroom climate show greater improvement on climate outcomes than interventions that indirectly target school climate or classroom climate (e.g. social emotional learning, anti-bullying, leadership and behaviour interventions that are theorised to indirectly improve climate)?**

These questions were viewed as important after feedback from EEF colleagues and our advisory group of academic and school representatives. Potential additional questions on effectiveness of intervention types and implementation are now main research questions as they were considered high priority. Some other questions on digital aspects of school and classroom climate, overlap between classroom climate and social emotional classrooms, and identifying measures of school and classroom climate that can be used by schools were considered less of a priority. Research question 6b fits well with RQ2 and will be reported alongside the effectiveness findings.

This protocol next moves on to specify the methods to be used for Review 2.

Methodology

This systematic review examines how interventions aimed at improving SC/CC affect pupil and teacher outcomes. Our proposed systematic review will follow best practice conduct and reporting guidelines (Akers et al., 2009; Aloe et al., 2024; Page et al., 2021), including extensions relating to equity (Tong et al., 2012).

Methodologically, Review 2 is a theory-driven, mixed methods review of SC/CC intervention evaluations. For the theory driven element, we will draw upon guidance from White (2018) and the example from Pierce et al. (2025) on conducting theory-based systematic reviews. The review is theory-driven in three key ways. First, we will assess whether included interventions align with the conceptualisation of school climate (SC) and classroom climate (CC) established through

evidence synthesis in Review 1. Second, we will build on this foundation to examine how, why, for whom, and under what conditions interventions are effective, further developing and refining logic models that set out the causal pathways linking contextual factors and intervention features to proximal and distal outcomes. Third, when analysing which combinations of intervention components, delivery features, and contextual conditions are associated with improvements in key outcomes (such as behaviour and attainment), we will draw on theoretically informed hypotheses derived from the literature to guide the selection and interpretation of relevant factors.

A mixed methods review fits well as we are interested in moving beyond only *whether* SC and CC interventions are effective and we therefore will not solely rely on quantitative evaluations of intervention effects. Qualitative research, such as process evaluations of SC/CC interventions, will help to explain intervention mechanisms, implementation factors, and contextual moderators. It will therefore help to interpret how, why and for whom causal pathways in our logic models apply. Furthermore, qualitative research can be useful in explaining moderation information about subgroups and contextual factors, such as differential engagement across groups, reported barriers or facilitators to participation, or authors' reflections on how contextual factors (e.g. school resources) shape intervention impacts. We will retain flexibility when conducting the review, in line with Pierce et al. (2025), so we can respond to important lines of enquiry during the review. For instance, we will also draw on the Review 1 search to explore overlap between SC/CC and other overlapping terms.

The focus of our theory-driven, mixed methods systematic review will be on primary research. Interventions may aim to improve SC/CC directly or indirectly, meaning it will be important to locate studies that identify SC/CC either in the intervention description or as a primary outcome predicted to improve for participants receiving an intervention. We will not restrict included interventions to those that only match our conceptualisation of SC/CC. This is because a) we want to explore what proportion of interventions do fit our conceptualisation drawn from the literature, b) schools are interested in improving SC/CC and this may be done indirectly through interventions with other immediate targets (e.g. behaviour or anti-bullying), and c) we will not require interventions to address all domains of SC or CC as identified in Review 1. For example, interventions that focus on a subset of school climate domains (e.g. relationships and safety) or selected classroom climate dimensions (e.g. student participation, teacher–student relationships, and behaviour management) will still be considered relevant. Rather than excluding such studies, we will treat variation in coverage as analytically informative. This approach allows us to compare interventions that align with our conceptualisation to differing degrees, and to explore whether more comprehensive interventions – those addressing multiple or all domains – are associated with stronger or more consistent effects (RQ4).

Inclusion and exclusion criteria for the review

This section sets out the eligibility criteria that will guide the inclusion and exclusion of studies in Review 2. These criteria ensure that the evidence selected is relevant, rigorous, and aligned with the review objectives, while maintaining transparency about the rationale for any exclusions.

Category	Criteria
Context/setting	<i>Include</i> studies primarily conducted in age 3-18 school settings. <i>Exclude</i> studies that are very different to the English education system: Pedagogical Approach <i>Exclude</i> studies from systems where rote learning and memorisation dominate classroom practice, leaving little scope for interaction or student voice. Teacher Authority and Role <i>Exclude</i> contexts where teacher authority is highly hierarchical and prescriptive, and teachers have minimal autonomy to shape CC. Resource Constraints <i>Exclude</i> studies from settings with severe resource limitations (e.g., very large class sizes, lack of basic materials) that fundamentally alter classroom dynamics. Cultural Norms <i>Exclude</i> systems where cultural expectations strongly discourage student participation, questioning, or collaborative learning.

	<i>Exclude</i> nursery childcare and Higher Education settings.
Participants	<i>Include</i> studies where intervention recipients are any members of a school or classroom community. This can include families. <i>Include</i> any participants who can provide outcomes related to the intervention, not only intervention recipients. This can include those who deliver the intervention or school leaders' views about classroom climate interventions they are not directly involved in.
Study design	<i>Include</i> study types that evaluate a SC/CC intervention. Quantitative evaluations must include a comparison group (e.g. usual practice, waitlist, matched comparison, alternative intervention). Therefore single group pre-post designs will be <i>excluded</i>. Process evaluations and qualitative research about eligible interventions will be <i>included</i> (these studies do not require a comparison group). <i>Included</i> studies do not need to be mixed methods. <i>Exclude</i> reviews, but we will use the included studies from relevant reviews as a source for additional searching. Reviews will not be <i>included</i> as studies to maintain comparability across <i>included</i> studies and because no review to date has used the synthesis methods we propose. <i>Exclude</i> other study types, including: Descriptions of interventions, studies that measure SC/CC without intervening, book reviews, conference abstracts, editorials and opinion papers.
Intervention	<i>Include</i> interventions that aim to improve SC/CC. Often <i>included</i> studies will describe the intervention as targeting SC/CC and also measure SC/CC. <i>Included</i> studies do not have to measure an SC/CC outcome (e.g. an intervention that clearly fits our conception of SC and that is described as aiming to improve SC may focus on a distal outcome only). <i>Included</i> studies may aim to improve SC/CC directly or indirectly. Indirect interventions focus on other related factors that are theorised to then improve SC/CC (e.g. an anti-bullying intervention can also target safety, relationships and improve the school environment or a relational approach can also target participation, psychological safety, behaviour management and involvement of the local community). The intervention must clearly explain how it indirectly aims to improve SC/CC to be <i>included</i>. <i>Exclude</i> interventions that are not aiming to improve SC across the whole school or CC across the whole class. For instance, a minority of studies will only target subgroups of pupils, aiming to improve their SC/CC perceptions only. <i>Include</i> where intervention recipients are not whole school/class but the whole school/class is impacted (e.g. teacher professional development where practice impacts whole classes taught). Interventions that aim to improve overlapping constructs that were shown in Review 1 to fit SC/CC will be <i>included</i>. These terms <i>include</i> classroom environment, school organisational climate, school environment, school connection, school ethos. However, each study will be assessed for its fit with SC/CC and very different or partial fits with our conceptualisation will be <i>excluded</i>. <i>Include</i> any intervention that labels itself SC/CC even if it does not fit our conceptualisation. <i>Include</i> any intervention that does not label itself SC/CC that does fit our conceptualisation. <i>Exclude</i> school culture and belonging interventions, unless they are clearly positioned as targeting SC (neither construct from Review 1 sufficiently overlaps with SC). <i>Exclude</i> interventions that only focus on skills development on aspects of domains, not developing SC/CC. For instance, not every relationships intervention will focus on the experience and perception of relationships and community across the school. Where there is doubt about whether an intervention is focused on SC/CC, <i>include</i> interventions that address and aim to improve at least two domains from Review 1 conceptualisation clearly (e.g. relational and behavioural experiences and perceptions in the classroom) or that reflect and aim to improve dimensions from across all Review 1 domains (e.g. participation + prosocial behaviour + student engagement + facilities across the school).

Outcomes	We will not use outcomes as an inclusion criterion. While we expect that the majority of <i>included</i> studies will measure SC/CC outcomes, we also expect that: Some studies will <i>include</i> interventions that clearly target SC/CC but measure another outcome (e.g. engagement, behaviour, attainment). Such studies should be <i>included</i> . Some studies will <i>include</i> SC/CC as a secondary outcome to explore if an intervention improves climate (or to assess whether climate is adversely affected) even though the intervention does not directly or indirectly aim to improve SC/CC. Such studies should be <i>excluded</i> .
Language	<i>Include</i> English language studies only.
Date	<i>No date restrictions</i>
Publication status	<i>No publication type restrictions. Include</i> theses and grey literature that meet other inclusion criteria.

The review is interested in 3-18 year old school settings, so includes early years foundation stage but not Higher Education or 0-2 years childcare. Synthesis must hold relevance to the English education system, so very different school contexts may be excluded. However, non-mainstream provision will be included as we are interested in how SC/CC intervention delivery and outcomes may vary across contexts and for different pupil groups. We do not insist that included interventions fit the conceptualisation of SC/CC we reached in Review 1, but we will analyse whether interventions fit the conceptualisation and whether this is associated with effects. We know from Review 1 that most research refers to SC/CC directly, but some studies may use other related terms (school/classroom environment, school organisational climate, school connection, school ethos in particular). Interventions must aim to improve SC/CC as indicated by study authors to be included or be assessed as fitting our Review 1 conceptualisation. Interventions can aim to improve SC/CC directly or indirectly and must target the whole school/class. Evaluations may be quantitative, qualitative or mixed methods. Where evaluations are focused only on quantitative outcomes, study designs must include a comparison group. Only English-language studies are included to ensure conceptual accuracy without depending on translations. As SC/CC has been considered in literature for 75+ years we would not wish to exclude any long-standing interventions through restrictions on date. We will include grey literature, including theses, as intervention evaluations may well be published outside of academic journals and relying on only peer reviewed journal articles may lead to publication bias. Quality appraisal will shape interpretation and synthesis, ensuring evidence strength is considered throughout but not as a reason to exclude studies (see below). In Review 1 we found very few studies (n=6) that considered school and classroom climate together, therefore we will include one set of studies targeting CC and another targeting SC, anticipating that a minority may fit both.

Search strategy for identification of studies

We will search the following bibliographic databases: PsycINFO, MEDLINE, Social Policy and Practice via Ovid; ERIC, British Education Index, Education Research Complete, CINAHL via EBSCOhost; Cochrane Database of Systematic Reviews; Campbell Systematic Reviews; Social Science Citation Index via Web of Science; Australian Education Index, ProQuest Dissertations & Theses Global, Education Database, Social Science Database via ProQuest, Trials Register of Promoting Health Interventions and Bibliomap.

We will use a search strategy that balances sensitivity with precision, designed by our experienced information specialist using a combination of relevant subject headings, when available, and free text terms. Sets of terms are included that relate to the following inclusion criteria:

- School climate terms (including environment, connection, ethos), e.g. school* climat*, school* ethos, school* connect*, school* environment*, organi* near climate
- Classroom climate terms (including environment), e.g. class* climat*, class* environment*
- Intervention terms, e.g. approach*, deliver*, interven*, program*

To be retrieved by the search, a term relating to either SC or CC needs to be used in title, abstract or subject heading, as well as a term relating to interventions. Compared to Review 1 we have removed terms related to engagement as this most often relates to outcome measures and are typically distinct aspects of climate. We will not search for school and classroom culture as this holds different meaning to climate (norms, values and expectations). We include ethos,

organisational climate and environment as these terms are often used synonymously with SC/CC. We search for intervention terms, rather than study design terms, because the full range of mixed methods studies that may evaluate SC/CC interventions will be diverse and not consistently labelled. We have restricted use of the subject heading “school environment” so the search does not return all research located in school settings.

A search strategy as used for the database ERIC is in Appendix 5.

We will conduct searches for grey literature from web searches, repositories of grey literature (e.g. The WHO Institutional Repository for Information Sharing, CORE), and websites (e.g. The National School Climate Center; ed.gov safe learning environments). Where necessary we will use the terms “school climate” and “classroom climate” to narrow these searches.

We will use supplementary search methods to identify any studies not captured by our sensitive database searches. We will review the reference lists of systematic reviews included in Review 1, review the reference lists of relevant reviews located in the search for Review 2, contact authors and experts, forward and backward citation search included studies and perform targeted searches using author names. We will hand-search any journals not indexed in databases searched that have published included studies.

Selection of studies

Search results will be deduplicated and uploaded to EPPI-Reviewer software. Reviewers will screen independently, assisted by EPPI-Reviewer’s priority screening functionality, assuming more than 6000 records to screen on title and abstract. Study selection will involve title and abstract screening and then full text screening in line with inclusion criteria as indicated above. This will be completed by two independent reviewers, with disagreements referred to a third reviewer. All reviewers will pilot a set of title and abstract and full text screening to align decisions and interpretation of inclusion criteria at the start of each screening stage. Inclusion criteria will be refined as necessary to ensure clarity and increase agreement, with any substantive changes reported. We will also use examples of borderline decisions during piloting to aid screening decisions and bring consistency to the process followed. Results of the screening process will be documented using a PRISMA-style flow chart.

Data extraction and management

Data extraction will focus on details of the interventions as well as the findings reported from evaluations of them. Two reviewers will independently extract data for all included studies using standardised, piloted forms, with discrepancies resolved through discussion or consultation with a third reviewer. The piloting will involve 10 diverse studies. For all studies, we will extract information on study details, study design and methods, intervention components and delivery, outcome measures and findings. Our data extraction will incorporate the EEF database main data extraction form to aid data sharing. We will also extract equity-relevant characteristics using PROGRESS-Plus (Karran et al., 2023) and protected characteristics where available, including country, ethnicity, age, sex and gender categorisation as reported, disability, socioeconomic status, children in care and children with a social worker. This will support RQ5.

Because Review 1 identified substantial variation in how interventions relate to SC and CC, we will not impose a fixed intervention taxonomy at the data extraction stage. Instead, we will extract detailed information on intervention aims, components, delivery approaches, target populations, target group (e.g. school-wide, classroom-level, targeted), intensity, and duration. This will allow intervention categories to be developed iteratively from the evidence base while remaining informed by the conceptual domains identified in Review 1. We anticipate interventions may include, for example, SC/CC interventions, social and emotional learning interventions, behaviour interventions, leadership and organisational interventions, anti-bullying interventions, and other approaches theorised to influence climate. Categories may be refined during synthesis where distinct interventions are frequently seen in the included literature base (e.g. if sufficient studies evaluate a specific approach such as School-Wide Positive Behavioural Interventions and Supports).

Data will be extracted for all relevant outcomes reported within included studies. During synthesis, outcomes will be grouped into broader outcome domains to support comparison across studies using different measures. Based on Review

1, we anticipate these domains may include climate outcomes (SC and CC), attainment, attendance, behaviour, bullying, engagement and motivation, wellbeing and mental health, and staff outcomes, although the final outcome framework will be informed by the outcomes identified in the included studies. We will extract information on the specific measures used, their conceptual alignment with our definitions of SC/CC, whether they are validated existing measures, and the timing of assessments. Outcome timing will be recorded as reported by study authors and subsequently organised into comparable follow-up periods for synthesis where feasible.

As indicated above it will be important to extract detailed components of interventions and SC/CC outcome measures to assess which interventions and outcomes fit our Review 1 conceptualisation and to ensure meta-analysis includes studies with similar interventions and outcomes. We will use EPPI-Reviewer or Microsoft Excel for this data extraction. Final data extraction will be uploaded to EPPI-Reviewer to aid data sharing.

Appraisal of included studies

We will assess and report study quality across all included studies. Two reviewers will be involved in the appraisal process: one reviewer will complete the quality assessment and a second reviewer will check the ratings. We will use Cochrane risk of bias tools for trials (Sterne et al., 2016; 2019), the EPPI-Centre tool for qualitative studies (Rees et al., 2011), and the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018) for studies that do not fit standard trial or qualitative designs.

Quality appraisal will be used to inform and guide the synthesis, rather than as a basis for excluding studies. Specifically, we will consider study quality when interpreting findings, including examining whether reported effect sizes vary according to study quality across different intervention types and outcomes. We will also highlight where evidence relating to implementation factors appears strong but may be supported by studies of lower methodological quality, ensuring that conclusions are appropriately cautious and transparent.

We will assess the certainty of evidence for all key findings using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach, considering the domains of study limitations (risk of bias), indirectness, inconsistency, imprecision, and publication bias (Guyatt et al., 2008). GRADE will be applied primarily to quantitative findings, with certainty ratings undertaken at the intervention–outcome level across studies, enabling conclusions to reflect not only whether effects are observed but the strength and reliability of the supporting evidence.

Qualitative and mixed-methods evidence relating to implementation processes, mechanisms, and contextual influences will be appraised separately using complementary approaches (e.g. GRADE-CERQual, Lewin et al., 2018). Bringing these strands together, overall certainty assessments will inform our conclusions by distinguishing between (a) interventions supported by robust, high-certainty evidence of effectiveness for specific outcomes, and (b) those where evidence is more limited, uncertain, or context-dependent.

Assessing certainty across both effectiveness and implementation evidence will support a more nuanced understanding of not only whether interventions work, but how and why they work, for whom, and under what conditions – an essential consideration in mixed methods evidence synthesis of complex school-based interventions.

Data synthesis

Five synthesis methods will be used to analyse included studies to answer the review questions and help conclude how, why and for whom SC/CC interventions are effective. Because narrative synthesis is used with different purposes according to research question, this section is organised by the synthesis to be conducted to answer each research question.

Initial descriptive synthesis (text and tables to provide an initial descriptive summary and explanation of the characteristics of the included studies, Akers et al., 2009) will report the key details about setting, aim, intervention, study design and samples of included studies.

RQ1: How are interventions targeting school and classroom climate designed and delivered? To what extent do they align with the conceptualisation of school and classroom climate established in Review 1?

We will use a narrative synthesis approach to describe the interventions across included studies and reporting on their fit with domains of SC/CC identified in Review 1 (1. Relational and social-emotional; 2. Teaching and learning; 3. Behavioural and organisational for CC | 1. Relationships and community; 2. Safety and behaviour; 3. Teaching and learning; 4. School environment for SC).

To robustly report on intervention components that are more fine-grained than category of intervention (SC versus anti-bullying) or the extent to which interventions target different domains (safety versus community), we will draw upon Intervention components analysis (ICA) to identify specific features of SC/CC interventions likely to be causally linked to contextual factors and outcomes (Sutcliffe & Kneale, 2025). ICA is an inductive approach to comprehensively describe and categorise intervention components which is appropriate when we cannot assume that interventions will clearly map to the domains and sub-domains of SC/CC identified in Review 1 (Sutcliffe et al., 2015). Because interventions aiming to improve SC/CC may vary widely in content, intensity, and delivery mode, examining them as whole packages risks obscuring which elements drive change. ICA will enable us to:

- Identify candidate active ingredients by inductively coding intervention features across studies, constructing a component-by-study matrix, and examining whether particular components or configurations of components occur more frequently in interventions associated with positive outcomes, (e.g., pupil voice, psychological safety, expectations).
- Distinguish components that function differently across contexts, such as age of pupils, culturally tailored materials, teacher-led vs. peer-led delivery.
- Improve explanatory power of causal chain analysis and qualitative comparative analysis (see RQ 4), ensuring intervention features used in analyses across studies are anchored in observable, replicable intervention components.

Incorporating ICA strengthens the overall synthesis by identifying specific, actionable programme features, which may be linked to implementation, school context and outcomes.

RQ2: How effective are different types of school and classroom climate interventions in improving pupil outcomes, including attainment and non-cognitive outcomes (e.g., wellbeing, engagement, behaviour), as well as school staff outcomes?

It will not be feasible or appropriate to provide overall effect sizes for all SC and CC interventions, given the expected heterogeneity in study designs, intervention targets, components, and outcome domains. Therefore, we will report and compare effects for different types of SC/CC interventions across different types of outcomes reported. Quantitative evaluations will be included in this synthesis and we will prioritise evidence from RCTs where available or non-randomised studies with control groups if <2 RCTs are available for any intervention + outcome dyad.

Studies will first be organised by intervention type, using categories informed by Review 1. Outcomes will be grouped into categories such as behaviour, engagement, attainment, rather than more fine-grained categories, given the wide range of outcomes seen in reviews synthesised in Review 1.

Where data permit, we will undertake meta-analysis using random effects models with robust variance estimation (RVE). RVE allows inclusion of all relevant effect sizes while adjusting standard errors for within study dependence, improving on alternative approaches such as selecting a single effect size, aggregating within studies, or implementing multilevel models when dependence structures are unknown (Zejnullahi & Hedges, 2024). This approach is increasingly used in education systematic reviews (e.g. Fisher et al., 2023).

Prior to undertaking each meta-analysis, we will assess whether enough RCTs are available for the relevant intervention category, outcome domain, and follow up period. Where at least two sufficiently homogeneous RCTs are identified, these studies will be synthesised separately and prioritised for meta-analysis. In such cases, non-randomised trials and quasi-experimental studies will be synthesised narratively with effect sizes reported where they can be calculated. Where insufficient RCTs are available, evidence from non-randomised and quasi-experimental studies will be considered for meta-analysis, subject to assessment of methodological comparability and necessary outcome data.

Continuous outcomes will be converted to standardised mean differences using appropriate transformations. Effect sizes will be calculated for all eligible studies, including non-RCTs, with Hedges' g used as the preferred standardized mean difference estimate. Effect sizes will be meta-analysed within intervention categories, outcome domains, and follow-up periods. All estimates will be based on intention to treat data where available (i.e. for RCTs).

Cluster randomised trials will be examined to ensure analyses have appropriately adjusted for clustering (Heo et al., 2018). We will extract intracluster correlation coefficients (ICCs), cluster sizes, and design effects, where reported. Where clustering has not been accounted for, we will adjust standard errors using estimated ICCs derived from comparable studies, with sensitivity analyses across plausible ICC values. Statistical heterogeneity will be examined using Cochran's Q , τ^2 , and the I^2 statistic.

The main output for RQ2 and meta-analyses run will be a matrix of pooled effect sizes and confidence intervals (see Appendix 6 for an example). This output will allow comparison of how effects vary across different types of interventions and outcomes. Colour coding of cells will be used to help show patterns where effects are stronger. The matrix will also support comparisons between interventions that directly target climate and those that aim to influence climate indirectly. These comparisons will be used to answer RQ6b: Do interventions that directly target school climate or classroom climate show greater improvement on climate outcomes than interventions that indirectly target school climate or classroom climate (e.g. social emotional learning, anti-bullying, leadership and behaviour interventions that are theorised to indirectly improve climate).

We will use narrative synthesis where meta-analysis is not possible due to fewer than two studies being sufficiently homogeneous in terms of intervention and outcome. Narrative synthesis will also be used when meta-analysis is used to ensure a transparent, structured, and reproducible synthesis (Campbell et al., 2020). It will be important to describe the similarities and differences amongst interventions and outcomes within each category. We will also comment on diversity in terms of intervention delivery, sample and setting where effects are pooled. This approach is particularly appropriate given the variability in outcome metrics observed in Review 1 and will support comparability while identifying patterns of effectiveness across intervention types and contexts.

RQ3: What conditions and strategies support the successful implementation and long-term sustainability of school and classroom climate interventions in diverse educational settings?

We will use narrative synthesis to report implementation evidence from included studies. This is appropriate as we anticipate mixed methods evidence in relation to implementation of SC/CC interventions. Where there are quantitative implementation outcomes these are likely to be focused on the SC/CC intervention evaluated rather than a comparative outcome. We also expect a range of qualitative outcomes to be relevant. We will draw on the EEF Implementation evidence review (Moore et al., 2024) for a framework to consider how implementation outcomes, implementation strategies and the contexts and behaviours that impact implementation may help to analyse what supports and challenges implementation of SC/CC interventions and whether this varies across intervention types or setting.

RQ4: What school climate and classroom climate intervention components and contextual factors (and combinations of these components and factors) are causally linked to improvements in key outcomes, like behaviour, mental health and attainment?

The previous syntheses for RQs 1-3 described above will help to identify and provide relevant data for our synthesis of RQ4. In particular, the ICA for RQ1 will identify intervention components, the work for any meta-analysis and narrative synthesis for RQ2 will identify frequently reported outcomes and calculate effect sizes for each study, the narrative synthesis for RQ3 will identify some relevant contextual factors related to implementation of SC/CC interventions. The syntheses for RQ4 aim to use this varied information about effectiveness, intervention components, implementation and contextual factors to go further than assessing which interventions are effective for particular outcomes and explore how intervention components and contextual factors can impact key outcomes.

Two main syntheses will be used to address this research question. Firstly, we will build on our causal chain analysis and logic modelling work in Review 1. We will extend the use of Causal Chain Analysis (CCA), which has been widely applied in theory-based systematic reviews (Kneale et al., 2018; White, 2018), to examine empirically observed causal pathways within intervention evaluation studies. Whereas Review 1 used CCA to construct mainly theoretically informed logic

models for SC and CC, Review 2 will use CCA to test, refine, and elaborate these models using primary intervention evidence. This approach enables a structured comparison between hypothesised pathways and observed outcomes, allowing us to identify which causal links are supported, modified, or contradicted by empirical findings. Because the logic models developed in Review 1 relied on theoretical work and systematic review findings, we also expect to add details currently missing from the logic models – for instance implementation factors and teacher outcomes for CC.

In Review 2, CCA will guide the systematic mapping of intervention evidence iterating the existing SC and CC logic models, enabling us to assess how intervention components and contextual factors contribute to outcomes in practice. We will explicitly examine where empirical evidence aligns with, strengthens, or challenges the causal assumptions identified in Review 1. This process will support a more robust and policy-relevant understanding of how SC/CC interventions operate, under what conditions they are effective, and for which populations.

To implement this approach, we will begin with the logic models developed in Review 1 as a priori frameworks, including the conceptualisations of SC/CC, hypothesised pathways, and identified contextual moderators (for SC in particular). We will then synthesise data which will already have been extracted from included intervention studies to populate, test, and refine these causal chains. Following guidance from White (2018) and Kneale et al. (2018), we will assess the extent to which intervention evidence demonstrates progression along the causal chain from inputs (intervention components), through intermediate outcomes (e.g. changes in relationships, engagement, or behaviour), to distal outcomes (e.g. wellbeing, attainment). Where studies report partial pathways, we will map these onto the relevant segments of the chain. This will mean that we can evidence causal chains involving factors, intervention components, proximal and distal outcomes by using multiple studies that provide smaller causal links.

Where evidence suggests divergence from the original logic model, we will iteratively refine the model to reflect observed rather than hypothesised causal pathways. This may include adding new proximal outcomes, modifying existing links, or differentiating pathways across school versus classroom levels. If, as expected, SC and CC operate through distinct mechanisms in practice, we will maintain separate, empirically grounded logic models.

This stage will also allow integration of quantitative evidence on effectiveness (e.g. effect sizes, direction of impact) with qualitative insights on implementation and intervention components, supporting a mixed methods interpretation of causal processes. The refined causal chains will therefore represent evidence-informed explanations of how interventions produce change, rather than solely theoretical propositions.

The steps for applying CCA in Review 2 will include:

1. Use Review 1 logic models as the starting framework
2. Gather extracted intervention evidence on components and outcomes
3. Map empirical findings onto existing causal chains
4. Assess strength and consistency of evidence for each link
5. Identify moderators and contextual influences
6. Compare observed (Review 2) vs hypothesised (Review 1) pathways
7. Revise and refine the logic models
8. Highlight gaps, contradictions, and uncertainties in the evidence base
9. Differentiate pathways where necessary (e.g. SC vs CC, phase differences)
10. Integrate quantitative and qualitative findings into coherent causal explanations

By extending CCA in this way, Review 2 moves from theory development to empirical testing and refinement, enabling us to understand not only whether SC/CC interventions are effective, but how causal processes operate in real-world settings, for whom, and under what conditions. This synthesis directly addresses Review 2's focus on intervention

effectiveness, components and factors that affect outcomes, building on and strengthening the conceptual foundations established in Review 1. As is typical for previous CCA, we do not anticipate that final logic models will indicate effect sizes for specific pathways, rather mixed methods synthesis will be narratively discussed to transparently indicate the strength of evidence for causal pathways in the model that do have sufficient evidence.

While CCA in this review provides a structured and transparent approach to mapping and refining hypothesised causal pathways, it is not likely to generate quantified estimates of the strength of relationships between intervention components and outcomes. Instead, CCA supports the identification and assessment of plausible causal links, alongside an appraisal of the strength of evidence supporting different parts of the causal chain. As such, it enables us to determine whether pathways are theoretically and empirically supported, but not to formally predict how combinations of factors jointly produce strength of effects given anticipated heterogeneity and lack of moderation analysis in included studies. To extend beyond this, and to examine how multiple intervention components, implementation features, and contextual conditions interact to shape effectiveness, we will employ Qualitative Comparative Analysis (QCA).

QCA is a case-based, configurational method developed by Ragin (2009) that is particularly well suited to the synthesis of complex interventions. In this review, each case will correspond to an individual intervention evaluation study with effectiveness outcomes. QCA enables the identification of non-linear and non-additive causal patterns, recognising that intervention effectiveness often arises from specific combinations of conditions rather than single factors acting independently. This makes it especially appropriate for SC and CC interventions, which are inherently multifaceted and implemented in diverse educational contexts.

QCA uses set-theoretic principles and Boolean logic to identify configurations of conditions that are associated with an outcome of interest (Moore et al., 2018). In this review, outcomes will include both academic attainment and non-cognitive outcomes (e.g., behaviour, engagement, wellbeing), depending on data availability across included studies. Conditions may include:

- Intervention components (e.g. relationships, safety, behaviour domains)
- Delivery features (e.g. whole-school vs classroom-level implementation, duration, training)
- Implementation factors (e.g. fidelity, staff engagement, leadership support)
- Contextual conditions (e.g. socioeconomic context, school phase, pupil characteristics)

A key strength of QCA is its ability to assess necessity and sufficiency. A necessary condition is one that must be present for an outcome to occur, whereas a sufficient condition is one that consistently leads to the outcome, although alternative sufficient pathways may also exist. QCA also accommodates equifinality, meaning that multiple different configurations may lead to the same outcome—an important consideration in SC/CC intervention research where different approaches may be effective under different conditions. It is recognised that QCA may not identify consistent pathways to effectiveness if the available evidence is sparse, heterogeneous, or characterised by contradictory configurations. In such cases, findings will be interpreted cautiously, and the absence of clear solution sets will be treated as a meaningful result, indicating that current evidence does not support generalisable configurations linking intervention features to outcomes. Running QCA for CC interventions will rely on enough diverse studies to explore configurations or may need a smaller number of predicted factors, assuming a smaller pool of CC studies.

QCA will be explicitly grounded in the logic models developed and refined through CCA in Review 2. These models will provide a theoretically informed basis for identifying candidate conditions and hypothesised relationships. In this way, QCA builds directly on prior synthesis work, moving from mapping and testing causal pathways (CCA) to identifying empirically supported configurations of conditions associated with effectiveness.

We will employ a fuzzy-set QCA (fsQCA) approach, which allows for degrees of membership in a condition or outcome (rather than binary presence/absence alone). Each condition and outcome will be calibrated onto a scale from 0 (full non-membership) to 1 (full membership), with intermediate values representing partial membership. For instance, interventions might be scored according to their coverage of the safety and behaviour SC domain from Review 1's conceptualisation. 0 would be absent, 1 would be fully reflected, and 0.33 might focus on only safety or behaviour.

Calibration will be informed by study data (e.g. effect sizes, implementation descriptions) and pre-specified. Calibration decisions will be piloted and checked by a second reviewer.

The analytic process will follow established guidance for QCA in systematic reviews (e.g. Thomas et al., 2014) and will involve the following stages:

1. Case and condition selection: Define cases (intervention studies) and identify a parsimonious set of conditions informed by the refined SC/CC logic models, theoretical relevance, data availability, and variation across studies.
2. Data table construction: Extract and code data for each case across all selected conditions and outcomes, including calibration into fuzzy sets.
3. Truth table construction: Generate a truth table representing all observed configurations of conditions and their associated outcomes.
4. Truth table assessment: Examine the distribution of cases across configurations, identify and resolve contradictory configurations, and set a consistency threshold (e.g. ≥ 0.8) for determining reliable relationships.
5. Boolean minimisation: Derive simplified solution sets that represent combinations of conditions associated with the outcome, using intermediate solutions informed by theoretical expectations.
6. Analysis of necessity and sufficiency: Assess which conditions or configurations are necessary or sufficient for effectiveness.
7. Consideration of logical remainders: Examine unobserved configurations and consider their implications for interpretation.
8. Interpretation: Interpret solution pathways in relation to the review questions, individual studies, and the SC/CC logic models.

Consistent with best practice, we will examine pathways both to effectiveness and to limited or no effectiveness, allowing us to identify not only enabling configurations but also combinations associated with weaker outcomes (Thomas et al., 2014). This dual analysis supports a more nuanced understanding of why some interventions show beneficial effects while others may not.

The final stage will involve exploring how we can integrate QCA findings within the refined CCA logic models to produce an evidence-informed account of how SC and CC interventions operate. While CCA will have identified and assessed individual pathways and components, QCA should provide insight into which combinations of components, conditions, and contexts are most consistently associated with positive outcomes. Together, these approaches will enable us to move from understanding whether and how interventions work to understanding which configurations work best, for whom, and under what circumstances. Where no stable or theoretically meaningful configurations are identified, we will report this transparently and use these findings to inform conclusions about the limits of the evidence base and priorities for future research.

RQ5: How do the effectiveness and implementation of school and classroom climate interventions vary across different school phases, settings, and populations – particularly in relation to age, socio-economic disadvantage, SEND, Children in Care and other forms of inequality?

To address RQ5, we will undertake a structured, equity-focused synthesis examining how the effectiveness and implementation of SC and CC interventions vary across different school phases, settings, and populations. Given the expected heterogeneity of interventions and the inappropriateness of a single pooled effect across all studies, we do not anticipate conducting meta-regression analyses. Instead, we will synthesise evidence using a combination of structured narrative synthesis and targeted subgroup analysis, supported where appropriate by descriptive quantitative summaries.

Studies that report subgroup analyses, comparative effectiveness designs, or process evaluations will be included in this synthesis. These may provide both quantitative evidence (e.g. differential effect sizes by subgroup) and qualitative or mixed-methods insights (e.g. variations in engagement, acceptability, or implementation across populations and contexts). We will examine whether intervention effects and implementation processes differ according to equity-relevant characteristics, including age/phase (e.g. primary vs secondary), socio-economic disadvantage, special educational needs and disabilities (SEND), children in care, and other forms of inequality. Where reported, these characteristics will be organised using the PROGRESS Plus framework, allowing a systematic consideration of social stratifiers and potential inequities.

Where quantitative subgroup data are available, we will extract information on the direction and relative magnitude of differential effects, noting whether interventions appear more or less effective for particular groups. However, given likely limitations in reporting and comparability, such findings will be synthesised narratively rather than statistically combined across studies. Qualitative and process evidence will play a key role in this synthesis by providing insight into how and why interventions may work differently across groups and contexts. For example, studies may report differential engagement among pupils from socio-economically disadvantaged backgrounds, variations in teacher capacity across school phases, or barriers to implementation in settings with higher levels of need.

To support transparent interpretation of patterns across studies, we will use structured tabulation and visual summaries (e.g. harvest plots or equivalent, Ogilvie et al., 2008) to present whether interventions appear to:

- Reduce inequalities (greater benefits for more disadvantaged groups)
- Maintain existing gradients (similar effects across groups)
- Widen inequalities (greater benefits for more advantaged groups)
- Show mixed or unclear patterns.

This approach is particularly appropriate for complex and heterogeneous evidence bases, where both quantitative and qualitative moderation information contribute to understanding variation in outcomes.

Findings from this synthesis will be interpreted alongside the causal pathways and configurations identified in earlier syntheses (meta-analysis, CCA and QCA), enabling us to assess not only whether SC/CC interventions are effective overall, but how their effectiveness and implementation are shaped by context, population, and inequality. We expect that often evidence of equity analysis will be limited or inconsistent; this will be reported explicitly, highlighting gaps in the literature and implications for future research and practice.

RQ6a: What distinguishes school and classroom climate from each other and the following, often overlapping, terms: school connection, belonging, school ethos, authoritative school climate?

A narrative synthesis will be used to compare key terms that are often used interchangeably with SC but which evidence from Review 1 suggests can be distinguished. Unlike the other syntheses, this analysis will draw on literature from both Review 1 and Review 2, because conceptualisation of these terms is as relevant to the RQ as the impact of interventions aligned to each construct. We will include a subset of studies for this synthesis, including some that are not eligible for other Review 1 or Review 2 syntheses. For instance, some studies state that school connection is distinct from SC. These studies are excluded from our syntheses conceptualising SC and exploring intervention effects, but will be very relevant to understand the similarities and difference between SC and school connection.

We will develop a structured comparative matrix to systematically map how SC and CC are defined in relation to each other and the set of overlapping constructs, school connection, belonging, school ethos, and authoritative school climate. Definitions identified in the literature will be coded and compared across several key dimensions: (a) scope (e.g. whole-school vs classroom-level focus), (b) level of analysis (individual perceptions vs collective/shared constructs), (c) theoretical framing (e.g. psychological, sociological, organisational, or policy-oriented), and (d) core attributes (e.g.

relationships, safety, teaching and learning practices, norms, values, participation). Particular attention will be paid to the relationship between SC and CC. In addition to comparing SC and CC with overlapping constructs, we will examine further whether the literature conceptualises SC and CC as distinct constructs operating at different ecological levels, overlapping constructs capturing similar domains within different settings, or hierarchical constructs in which CC is considered a component, antecedent, or manifestation of broader SC. We will also further examine whether similar domains (e.g. relationships and community versus relational and social-emotional) and apparently identical domains (e.g. teaching and learning) are similar enough to suggest consistent domains, and therefore whether SC constructs like safety and resources should feature in the conceptualisation of CC. The intervention literature will add to the small amount of conceptual literature that compared SC and CC in Review 1 to achieve this.

This structured comparison will allow us to identify which features are consistently associated with SC and CC, and which are more characteristic of related constructs. For example, we will examine whether concepts such as *belonging* and *school connection* are more narrowly defined as individual-level experiences, compared to SC/CC which are typically conceptualised as multi-dimensional, shared properties of the school or classroom environment. Similarly, we will explore how *school ethos* (often framed in terms of values, culture, and institutional identity) and *authoritative school climate* (focused on the balance of support and structure) align with, overlap, or diverge from broader SC/CC definitions.

Through this process, we will identify:

- **Conceptual overlap:** elements that recur across SC/CC and related constructs (e.g. relationships, support)
- **Distinctive features:** characteristics that differentiate SC/CC as broader, multidimensional constructs (e.g. inclusion of organisational structures, teaching and learning, and behaviour)
- **Relationships between constructs:** For instance, whether CC is conceptualised as a component, determinant, manifestation, or distinct but related construct within broader SC; whether school-level and classroom-level climates are expected to influence one another. Also, how SC/CC or other constructs are predicted to be outcomes affected by each other and where terms like belonging and school connection might be a component of SC as a broader construct.

We will also examine how these alternative constructs may vary in their use across school phases (e.g. primary vs secondary), settings, and policy contexts, including their relevance to the English education system. This will help determine whether distinctions between SC, CC, and related constructs are context-dependent. By explicitly mapping boundaries between SC/CC and overlapping concepts, this step will allow us to clarify what is (and is not) captured by SC/CC, ensuring conceptual precision for the other syntheses. Finally, we will recommend whether understanding SC/CC requires distinguishing them from related constructs and whether SC and CC should be regarded as separate but related constructs, different levels of the same construct, or conceptually interchangeable terms. This will help inform the final conceptualisation of SC/CC to be shared beyond the evidence review.

RQ6b: Do interventions that directly target school climate or classroom climate show greater improvement on climate outcomes than interventions that indirectly target school climate or classroom climate (e.g. social emotional learning, anti-bullying, leadership and behaviour interventions that are theorised to indirectly improve climate).

Findings from the RQ2 effectiveness synthesis will be used to compare interventions that aim to improve SC/CC directly with interventions that aim to influence climate indirectly. We will use intervention categories identified in Review 1 to categorise the indirect interventions, while remaining open to other groups of similar interventions (e.g. mindfulness). Depending on the diversity of the interventions we may further categorise direct interventions in terms of which SC or CC domains they are directly targeting. Where meta-analysis is feasible as described above for the synthesis of RQ2, differences in pooled effects for climate outcomes will be compared across intervention types, direct versus indirect at the broadest level and by intervention categories, such as relationships focused, safety focused, SEL, anti-bullying. Where meta-analysis is not possible, narrative synthesis will be used.

Reporting

The evidence review final report will present methods and findings for both reviews in full. The final report will be structured in a similar way to this protocol with overall background and rationale and evidence review objectives. It will then present methods and finding for each review in turn. The final report will have a discussion and conclusion chapter which will bring together key messages from across the reviews that indicate overall judgements on the quality (or certainty) of the evidence base for a finding or recommendation.

Registration

Once finalised, this protocol will be published on the EEF website and the International Database of Education Systematic Reviews (IDESR). We will maintain a record of any amendments to the protocol that are made and either update this version with significant changes or report deviations from protocol in the final report.

Timeline

A high-level timeline is included below with a more detailed timeline provided in Appendix 7.

Table 1: Timeline

Dates	Activity	Staff responsible / leading
June – July 2026	Publish Protocol	DM
June – July 2026	Run Search	AB/DM
July – August 2026	Study selection	DM and research team
August 2026	Data extraction and quality appraisal	DM and research team
September 2026	Descriptive synthesis	DM and research team
August - September 2026	Intervention synthesis (RQ1)	DM/YZ
September - October 2026	Meta-analyses and intervention effects (RQ2)	DM/TR
September - October 2026	Implementation synthesis (RQ3)	SBC/JS
October - November 2026	Causal Chain Analysis (RQ4)	DM/YZ
December 2026	Qualitative Comparative Analysis (RQ4)	DM/SBC
January 2027	Final report write up	DM and research team

Team

The Review 2 team and their roles are listed below. All members are from the University of Exeter.

Darren Moore – Project oversight, supervise researchers, methods design and analysis across reviews. Lead report writing with editing support from all

Eleni Dimitrellou – Systematic review methods, Quality assurance, Interpretation of findings, socioemotional learning and belonging expertise

Alison Bethel – Information Specialist. Search strategy design and execution, systematic review methods

Yuanya Zhang – Researcher: study selection, data extraction, quality appraisal, synthesis.

Josie Scammell – Researcher: study selection, data extraction, quality appraisal, synthesis.

Simon Benham-Clarke – Researcher: study selection, data extraction, quality appraisal, synthesis.

Katherine Gulliver – Researcher: study selection, data extraction, quality appraisal, synthesis.

Tom Ralph – disadvantaged pupils expertise, meta-analysis support, school engagement, knowledge translation, interpretation of findings

Neil Harrison – Trauma-Informed Expertise, interpretation of findings

We have established an advisory group to guide the design and interpretation of the review. This group brings together experts in SC research, evidence synthesis methods, and practical implementation in schools. It also includes education leaders and a youth advisory panel to ensure that findings are relevant, feasible, and informed by real-world perspectives. We will engage group members flexibly and therefore are unlikely to hold meetings of the whole, diverse group together. We will share emerging insights with the advisory group throughout the project to help refine the logic model and identify the most promising approaches for improving SC and CC in the English educational context.

Conflicts of interest

The work described in this protocol is being undertaken by researchers at the University of Exeter and funded by the EEF. The views expressed are those of the authors and not necessarily those of the EEF. There are no conflicts of interest to report.

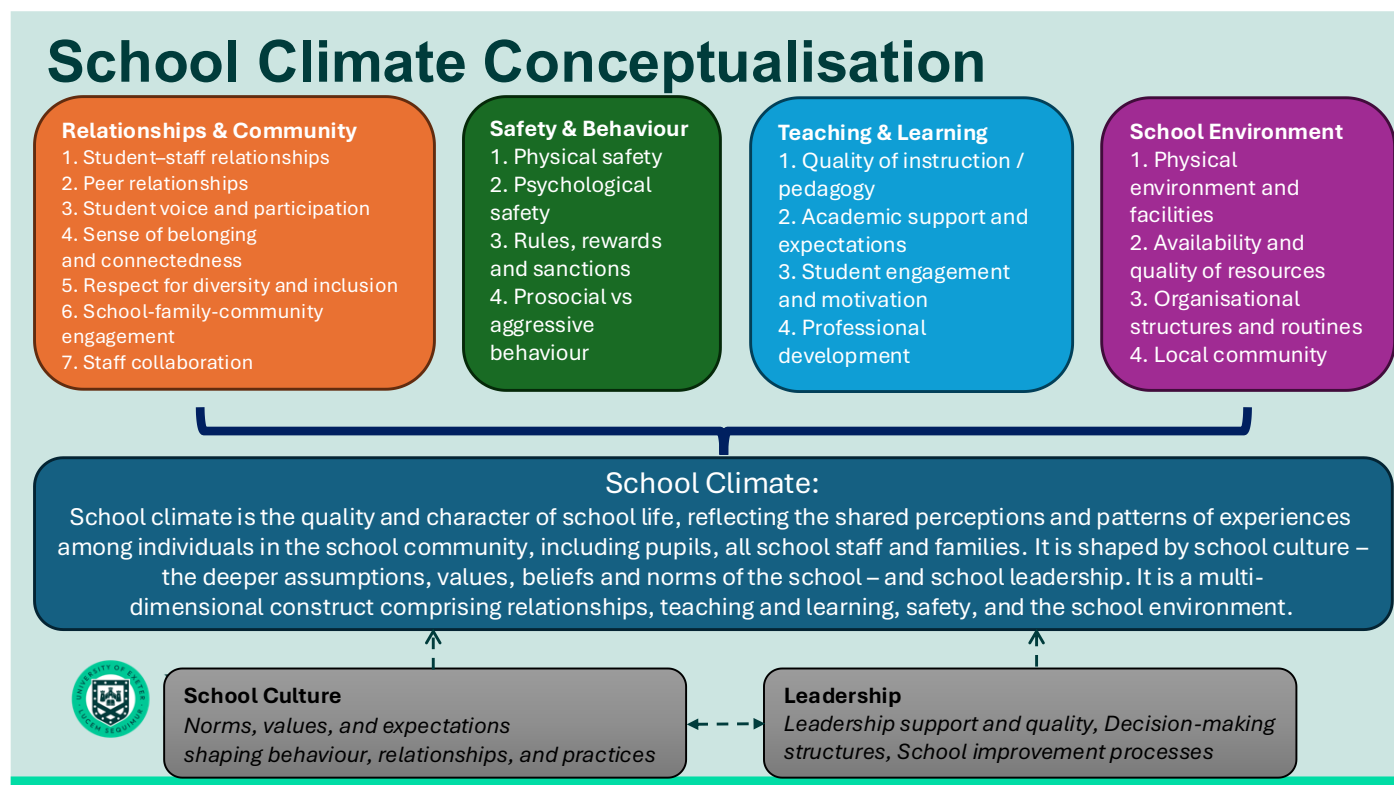
References

- Akers, J., Aguiar-Ibáñez, R., & Baba-Akbari, A. (2009). *Systematic reviews: CRD's guidance for undertaking reviews in health care*. Centre for Reviews and Dissemination, University of York.
- Aloe, A. M., Dewidar, O., Hennessy, E. A., Pigott, T., Stewart, G., Welch, V., Wilson, D. B., & Campbell MECCIR Working Group. (2024). Campbell standards: Modernizing Campbell's methodologic expectations for Campbell Collaboration intervention reviews (MECCIR). *Campbell Systematic Reviews*, 20(4), e1445. <https://doi.org/10.1002/cl2.1445>
- Baumsteiger, R., Hoffmann, J. D., Seibyl, J., Rose, B., & Brackett, M. A. (2023). A systematic review of secondary school climate assessments. *Educational Psychology Review*, 35(2), 47. <https://doi.org/10.1007/s10648-023-09748-y>
- Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., ... & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *bmj*, 368.
- Fisher, B. W., Petrosino, A., Sutherland, H., Guckenburg, S., Fronius, T., Benitez, I., & Earl, K. (2023). School-based law enforcement strategies to reduce crime, increase perceptions of safety, and improve learning outcomes in primary and secondary schools: A systematic review. *Campbell systematic reviews*, 19(4), e1360.
- Gillett-Swan, J., Monta, C. C., & Shanks, R. (2025). A systematic review of definitions of rights education for children: Children's rights education (CRE), children's human rights education (CHRE), and other rights education subtypes. *Review of Education*, 13(3), e70104.
- Grazia, V., & Molinari, L. (2021). School climate multidimensionality and measurement: A systematic literature review. *Research Papers in Education*, 36(5), 561-587. <https://doi.org/10.1080/02671522.2019.1697735>
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*, 336(7650), 924-926.
- Heo, M., Jimenez, C. C., Lim, J., Isasi, C. R., Blank, A. E., Lounsbury, D. W., ... & Wylie-Rosett, J. (2018). Effective nationwide school-based participatory extramural program on adolescent body mass index, health knowledge and behaviors. *BMC pediatrics*, 18(1), 7.
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., ... & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for information*, 34(4), 285-291.
- Karran, E. L., Cashin, A. G., Barker, T., Boyd, M. A., Chiarotto, A., Dewidar, O., ... & Moseley, G. L. (2023). Using PROGRESS-plus to identify current approaches to the collection and reporting of equity-relevant data: a scoping review. *Journal of Clinical Epidemiology*, 163, 70-78.
- Kneale, D., Thomas, J., Bangpan, M., Waddington, H., & Gough, D. (2018). Causal chain analysis in systematic reviews of international development interventions.
- Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., ... & Noyes, J. (2018). Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implementation Science*, 13(Suppl 1), 2.
- Moore D, Proctor R, Benham-Clarke S, Gains H, Melendez-Torres GJ, Axford N, Rogers M, Anderson R, Hall D, Hawkins J, Berry V, Forbes C, Lloyd J. (2024). *Review of evidence on implementation in education*. Education Endowment Foundation.
- Moore, D. A., Russell, A. E., Matthews, J., Ford, T. J., Rogers, M., Ukoumunne, O. C., ... & Gwernan-Jones, R. (2018). School-based interventions for attention-deficit/hyperactivity disorder: A systematic review with multiple synthesis methods. *Review of Education*, 6(3), 209-263.
- Ogilvie, D., Fayter, D., Petticrew, M., Sowden, A., Thomas, S., Whitehead, M., & Worthy, G. (2008). The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC medical research methodology*, 8(1), 8.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., & Chou, R. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pierce, M., Bai, Y., Taxiarchi, V., & Hugh-Jones, S. (2025). *Understanding drivers of recent trends in young people's mental health – Technical Report* (Technical Report). University of Manchester. https://youthfuturesfoundation.org/wp-content/uploads/2025/07/Technical-Report-Understanding-drivers-of-recent-trends-in-young-peoples-mental-health_University-of-Manchester_UCL_July-2025.pdf
- Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago press.

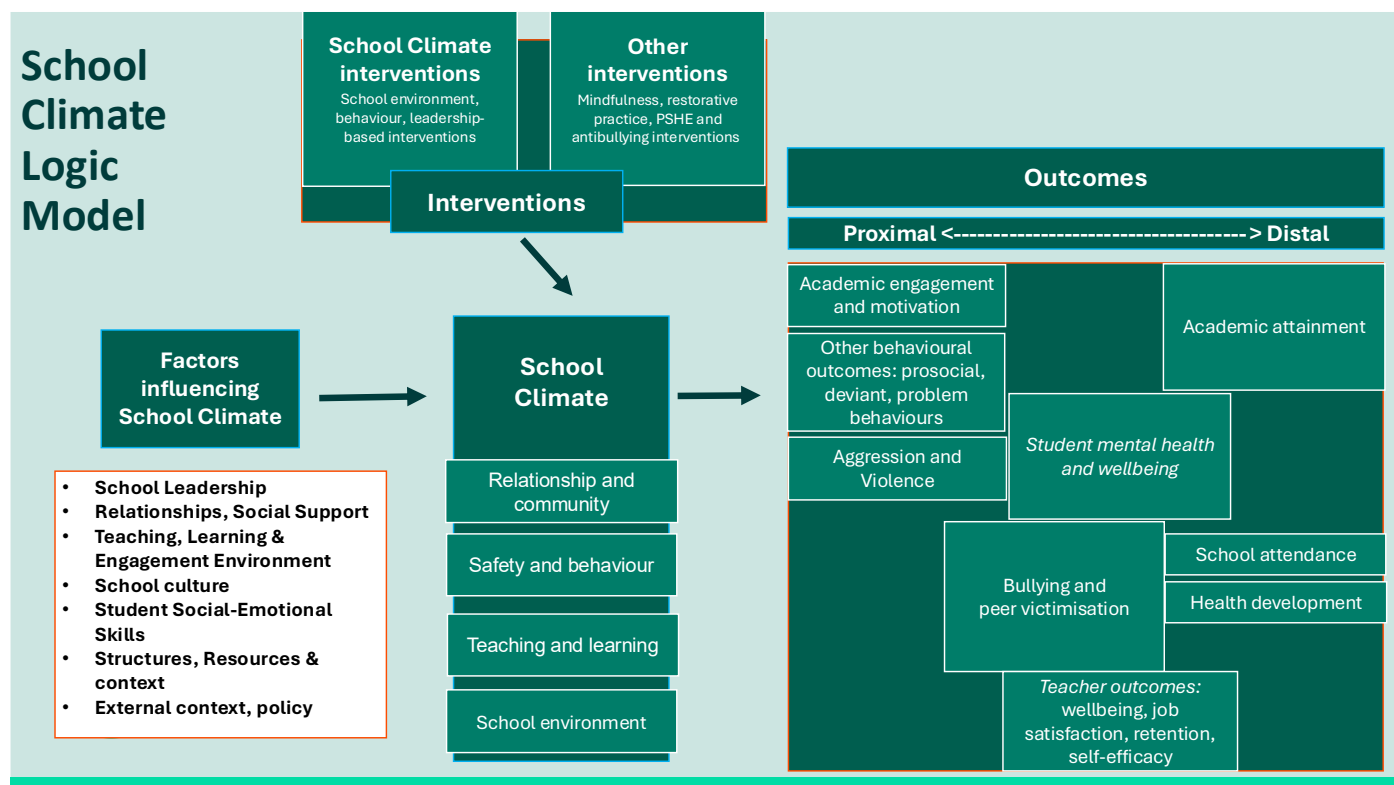
- Rees, R., Oliver, K., Woodman, J., & Thomas, J. (2011). The views of young children in the UK about obesity, body size, shape and weight: A systematic review. *BMC Public Health*, 11(1), 188. <https://doi.org/10.1186/1471-2458-11-188>
- Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., ... & Higgins, J. P. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *bmj*, 355.
- Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., ... & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *bmj*, 366.
- Sutcliffe, K., & Kneale, D. (2025). Identifying critical intervention, contextual and implementation features in systematic reviews: Intervention component analysis ten years on. *International journal of nursing studies*, 170, 105144.
- Sutcliffe, K., Thomas, J., Stokes, G., Hinds, K., & Bangpan, M. (2015). Intervention Component Analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. *Systematic reviews*, 4(1), 140.
- Thomas, J., O'Mara-Eves, A., & Brunton, G. (2014). Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. *Systematic reviews*, 3(1), 67.
- Tong, A., Flemming, K., McInnes, E., Oliver, S., & Craig, J. (2012). Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. *BMC Medical Research Methodology*, 12, 181. <https://doi.org/10.1186/1471-2288-12-181>
- White, H. (2018). Theory-based systematic reviews. *Journal of Development Effectiveness*, 10(1), 17–38. <https://doi.org/10.1080/19439342.2018.1439078>
- Zejnnullahi, R., & Hedges, L. V. (2024). Robust variance estimation in small meta-analysis with the standardized mean difference. *Research synthesis methods*, 15(1), 44-60.

Appendices

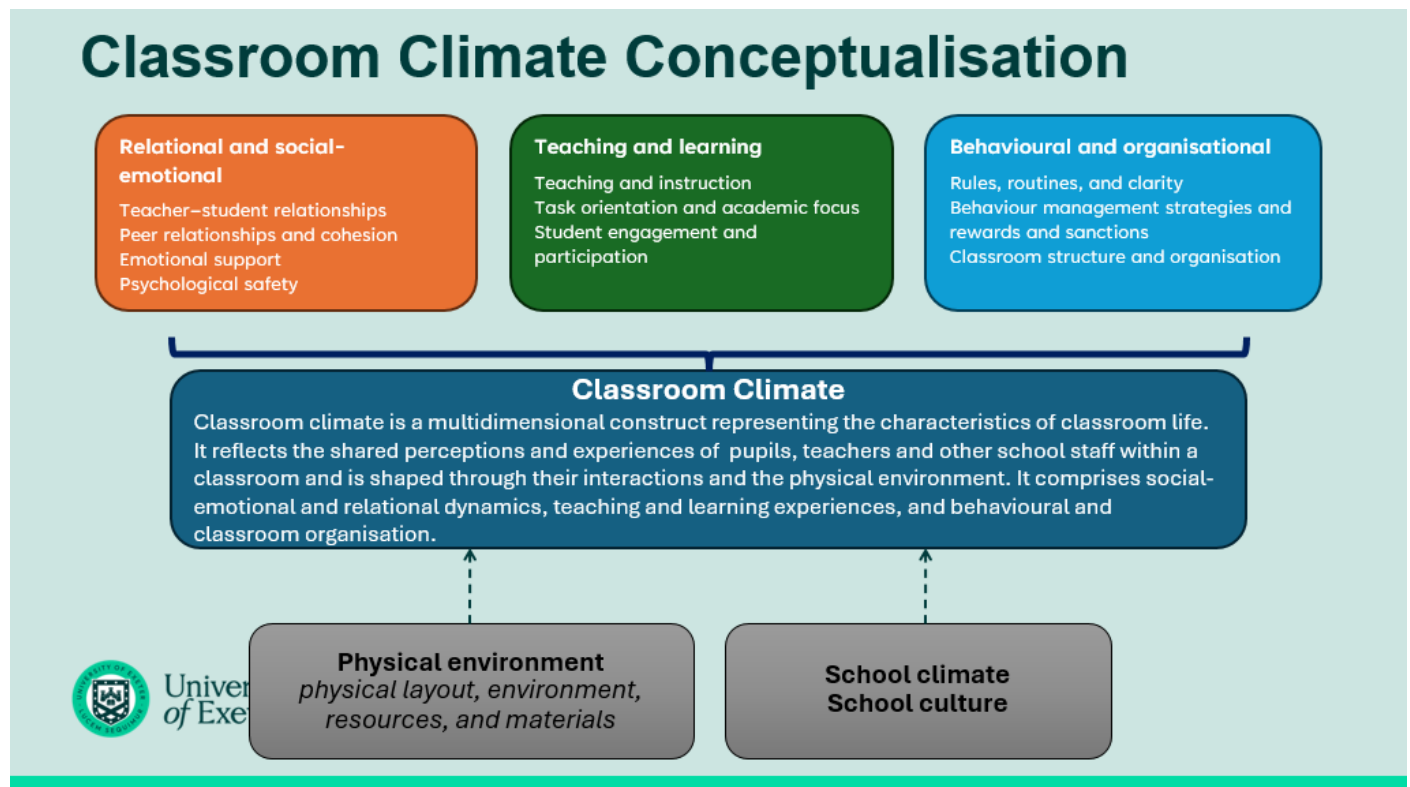
Appendix 1: Review 1 School climate conceptualisation



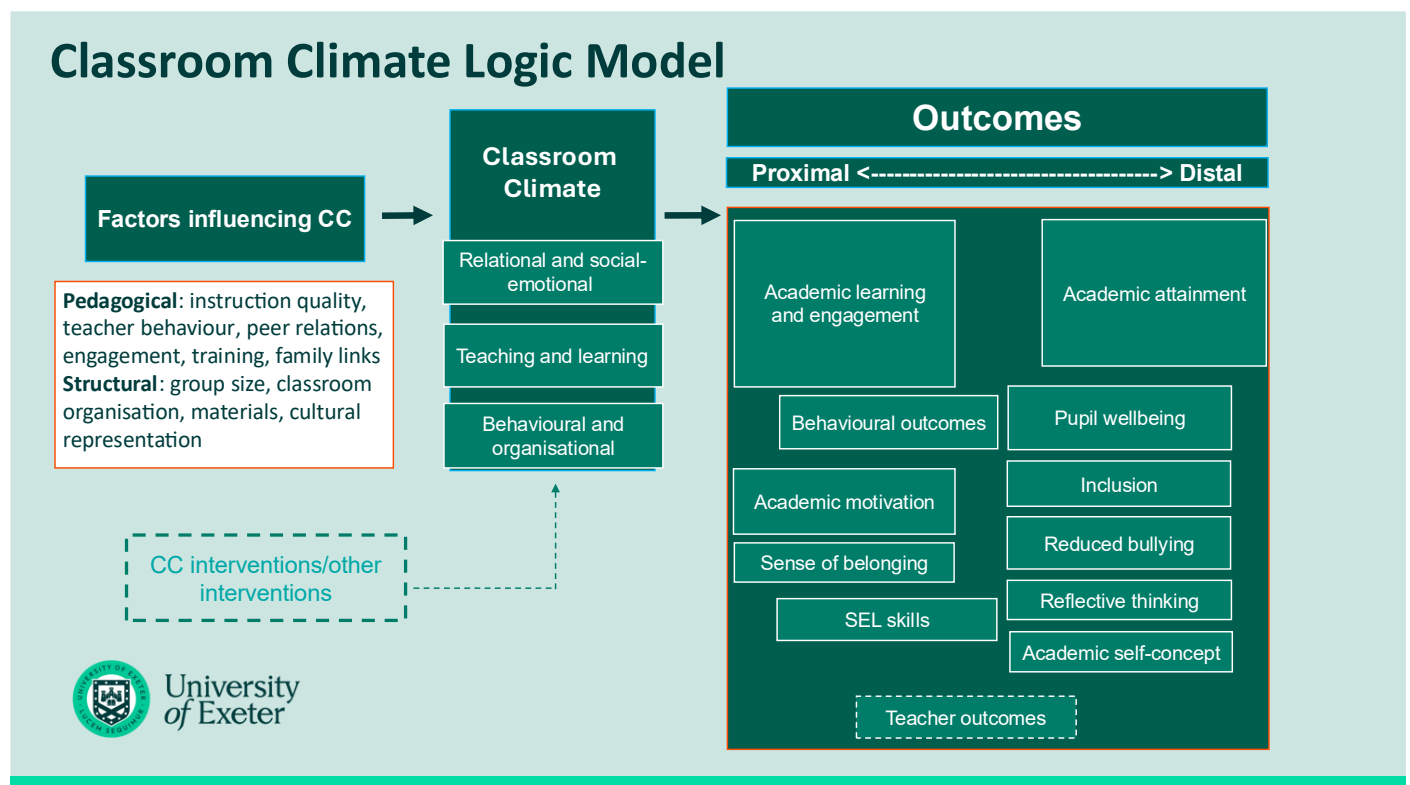
Appendix 2: Review 1 School climate logic model



Appendix 3: Review 1 classroom climate conceptualisation



Appendix 4: Review 1 Classroom climate logic model



Appendix 5: Example Search Strategy (ERIC)

ERIC search strategy draft 2. 29th June 2026

#	Query	Results
S17	S8 OR S16	3,434
S16	S14 AND S15	1,410
S15	S10 OR S12 OR S13	9,253
S14	TI (((belong* or connect* or climate* or ethos or environment) N3 (Approach* or curricul* or deliver* or evaluat* or framework* or implement* or initiative* or intervent* or policy* or policies or prevent* or program* or promot* or scheme* or strateg* or therap* or training))) OR AB (((belong* or connect* or climate* or ethos) N3 (Approach* or curricul* or deliver* or evaluat* or framework* or implement* or initiative* or intervent* or policy* or policies or prevent* or program* or promot* or scheme* or strateg* or therap* or training)))	11,575
S13	AB ((class or classes or classroom* or kinderg* preschool* or pre school* or school*) N3 (climate* or ethos or environment))	8,578
S12	S5 AND S11	766
S11	AB (organi* N3 (climate*))	1,435
S10	S2 AND S9	748
S9	AB ((class or classes or classroom* or school* or student* or pupil*) N3 (belong* or connect*))	16,610
S8	S3 OR S6 OR S7	2,430
S7	TI ((class or classes or classroom* or kinderg* preschool* or pre school* or	2,184

	school*) N2 (climate* or ethos or environment))	
S6	S4 AND S5	191
S5	TI ((class or classes or classroom* or kinderg* preschool* or pre school* or school*)) OR AB ((class or classes or classroom* or kinderg* preschool* or pre school* or school*))	841,635
S4	TI (organi* N2 (climate*))	348
S3	S1 AND S2	147
S2	TI (climate* or ethos or environment) OR AB (climate* or ethos or environment)	26,512
S1	TI ((class or classes or classroom* or school* or student* or pupil*) N2 (belong* or connect*))	2,614

Appendix 6: Snippet of example matrix of pooled effect sizes and confidence intervals

Intervention/ Outcome	School Climate Total	School Climate: Relationships	Bullying incidents	SEL competence
School Climate	0.6 (0.5-0.7)	0.5 (0.4-0.6)	0.3 (0-0.6)	0.1 (0-0.2)
Relationship domain	0.4 (0.3-0.5)	0.8 (0.7-0.9)	0.2 (0-0.4)	0.3 (0.2-0.4)
Anti-bullying	0.2 (0-0.4)	0.2 (0.1-0.3)	0.7 (0.6-0.8)	0.1 (-0.1-0.3)
Social emotional learning	0.3 (0-0.6)	0.4 (0.3-0.5)	0.1 (-0.1-0.3)	0.8 (0.6-1)

Appendix 7: Detailed Review 2 Timeline

Timeline									2027				
		June	July	August	September	October	November	December	January	February	March	April	May
	Set-up												
	Recruit Researcher/s												
	Advisory Group Meetings	■	■			■	■						
	EEF/Exeter Meetings	■	■		■	■	■	■	■	■			
Review 2 - Impact	Protocol (amends previous)	■											
	Searching	■											
	Study selection		■										
	Data extraction			■									
	Critical Appraisal			■									
	Descriptive Synthesis				■								
	Intervention Synthesis			■									
	Effectiveness synthesis				■								
	Implementation synthesis				■								
	CCA					■							
	QCA						■						
		Final Report write up							■			■	
		Other Dissemination								■		■	

You may re-use this document/publication (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0.

To view this licence, visit <https://nationalarchives.gov.uk/doc/open-government-licence/version/3> or email: psi@nationalarchives.gsi.gov.uk

Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned. The views expressed in this report are the authors' and do not necessarily reflect those of the Department for Education.

This document is available for download at <https://educationendowmentfoundation.org.uk>



**Education
Endowment
Foundation**

The Education Endowment Foundation
5th Floor, Millbank Tower
21–24 Millbank
London
SW1P 4QP

<https://educationendowmentfoundation.org.uk>



@EducEndowFoundn



Facebook.com/EducEndowFoundn