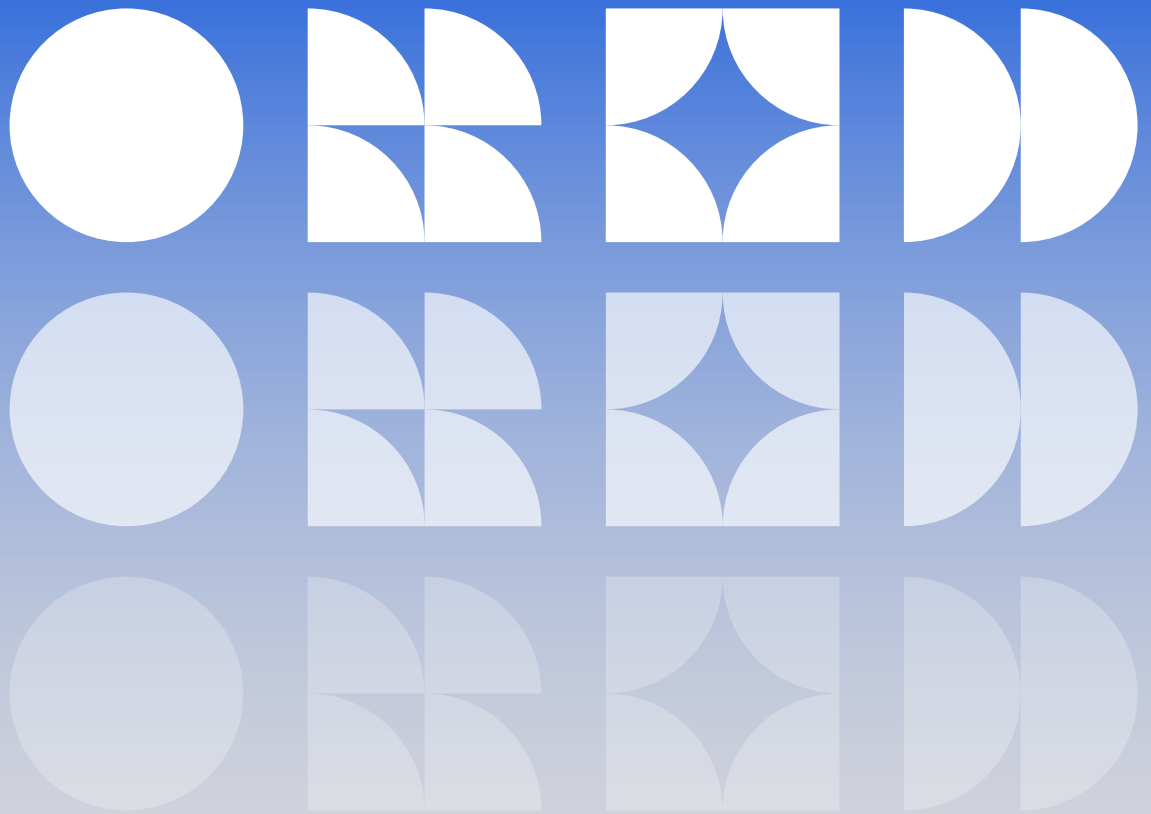




The Centre for Collaboration in Community Connectedness: Framework for Evidence Review & Synthesis

September 2026

Sum Yue (Jessica) Ko, Jie Qi, Ian Wilson



The Centre for Collaboration
in Community Connectedness



Economic
and Social
Research Council

The Centre for Collaboration in Community Connectedness: Framework for Evidence Review & Synthesis

Authors

Sum Yue (Jessica) Ko, Jie Qi & Ian Wilson (Sheffield Hallam University).

Funding

The Centre for Collaboration in Community Connectedness is funded by UK Research and Innovation (UKRI) as part of its work to create opportunities and improve outcomes.

Citation

Ko, S.Y. (J). Qi, J. Wilson, I. (2026). The Centre for Collaboration in Community Connectedness: Framework for Evidence Review & Synthesis. The Centre for Collaboration in Community Connectedness. <https://doi.org/10.7190/c4.2026.5271109646>

Version History

This is a living document, it will be reviewed and edited as appropriate.

Version	Date	Changes
1.0	September 2026	Not Applicable



**The Centre for Collaboration
in Community Connectedness**



**Economic
and Social
Research Council**

Contents

Page 4	1.	Introduction
4	1.1.	Background
4	1.2.	Purpose of this framework document
Page 5	2.	Developing our approach
5	2.1.	Scoping study on the nature of the evidence base on community connectedness
6	2.2.	Scoping study of the definition of community connectedness
6	2.3.	Desk review of existing What Works Centres
7	2.4.	Stakeholder interviews with evidence reviewers
9	2.5.	Stakeholder interviews with evidence users
Page 11	3.	Overview of our approach
11	3.1.	Guiding principles of our approach
12	3.2.	Stages of reviews
12	3.2.1.	Formulating the review question
12	3.2.2.	Registering protocols
13	3.2.3.	Search and retrieval of studies
14	3.2.4.	Study screening and selection
14	3.2.5.	Data extraction
16	3.2.6.	Quality appraisal and confidence assessment
17	3.2.7.	Synthesis
18	3.3.	Selecting review approach types
19	3.3.1.	Clarifying review approach types
20	3.4.	Engagement and co-production
20	3.5.	Use of artificial intelligence (AI) within reviews
20	3.6.	Report structure and output types
Page 22		References

1. Introduction

1.1. Background

The Centre for Collaboration in Community Connectedness (C4)

The Centre for Collaboration in Community Connectedness (C4) is a UK-wide research centre that brings together researchers, practitioners, policymakers, funders, and residents to improve understanding of how stronger, more connected communities can be developed and sustained. Established through support from the Economic and Social Research Council (ESRC) within UK Research and Innovation (UKRI), the Centre seeks to transform policy and practice relating to community connectedness by identifying what contributes to stronger relationships, greater participation, increased resilience, and enhanced community connectedness across the UK.

Evidence Review and Synthesis Unit within C4 (Implementation Unit 1/ C4 IU1)

As part of this overarching mission, C4's Implementation Unit 1 (C4 IU1), based within Sheffield Hallam University's Centre for Regional Economic and Social Research (CRESR), has been established to fulfil C4's evidence synthesis and review function. To this end, C4 IU1 is responsible for reviewing, synthesising, and translating evidence relevant to community connectedness through rigorous and transparent approaches. Drawing on the principles of the What Works movement while recognising the complexity of community systems, C4 IU1 seeks to not only understand whether interventions pertaining to community connectedness work, but also why they work, for whom, and under what circumstances and contexts.

1.2. Purpose of this framework document

This is a living document that establishes C4 IU1's current principles and approach to evidence review and synthesis, informed by methodological literature, scoping work and engagement with evidence producers and users. It is deliberately designed to evolve as C4 undertakes reviews across different questions, evidence bases and contexts. Experience from completed reviews will be used systematically to test, refine and extend the framework, including decisions about review design, searching, screening, appraisal, synthesis, stakeholder involvement, use of AI, reporting and evidence mobilisation. Major methodological changes will be documented through version control so that the development of C4's approach remains transparent.

As a living framework, this document should therefore be read as C4 IU1's current methodological position rather than as a fixed or exhaustive manual. Following completion of reviews, the review team will reflect on what worked well, what limitations or practical challenges were encountered, and whether methodological choices should be retained, adapted or developed further. Learning with implications beyond an individual review will be incorporated into subsequent versions of the framework. Where appropriate, review-specific protocols and reports will document methodological decisions that supplement or depart from the current framework.

The purpose of this framework document is threefold.

First, the framework sets out the principles, standards, and processes that guide the conduct of evidence reviews and evidence synthesis within C4 IU1. It is intended to **support IU1 researchers** in making clear and defensible decisions throughout the review process, from question development and protocol design through to synthesis, interpretation, and dissemination. In doing so, it seeks to minimise methodological inconsistency, reduce the risk of bias and post-hoc decision-making, and ensure that review findings can withstand scrutiny from academic,

policy, and community audiences.

Second, for C4 partners and stakeholders seeking to **collaborate on or commission** an evidence review, the framework provides guidance on C4 IU1's considerations involved in formulating review questions, selecting appropriate review methodologies, and determining the scope and intended use of review findings. It is intended to clarify what different review approaches can and cannot deliver, the resources required to conduct them, and the types of evidence that will be used to answer the review questions of partners and stakeholders, supporting the development of realistic expectations.

Finally, the framework, as a reference document, seeks to support community partners wishing to undertake their own evidence reviews and syntheses. While not intended to replace detailed methodological handbooks (such as the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2024)), it provides an accessible, simplified overview of what evidence synthesis entails, covering stages from protocol development and searching to synthesis, stakeholder involvement and dissemination of output. In doing so, it contributes to building evidence-review capability within and beyond the Centre and supports a wider culture of evidence-informed practice.

2. Developing our approach

Section two of this framework document details the various activities undertaken by C4 IU1 in developing our approach to systematic reviews and evidence syntheses for topics on community connectedness. The lessons learned from these activities have informed our approach in both direct and indirect ways.

2.1. Scoping study on the nature of the evidence base on community connectedness

What we did:

An initial scoping study mapped the nature of evidence relevant to community connectedness across disciplines and sectors. The study considered variation in geographical and institutional contexts, target populations, problem areas, theoretical foundations, proposed mechanisms, intervention modes, outcomes and evaluation methods. It also identified recurring problems reported across existing reviews, including inconsistent terminology and measures, weak distinction between intervention aims, mechanisms and outcomes, and the grouping of heterogeneous interventions despite differences in theoretical basis or delivery mode.

What we learned:

The study indicated that the evidence base is fragmented and methodologically diverse rather than organised around a single intervention tradition or outcome framework. Relevant evidence spans urban planning, environmental psychology, community development, social policy, public health and education, and uses experimental, quasi-experimental, qualitative, participatory, systems and spatial approaches. This limits the usefulness of organising evidence only by intervention label or sector. The mapping instead highlighted the importance of retaining distinctions between context, population, mechanism, intervention, outcome and evaluation design while still enabling comparison across domains. By disaggregating the evidence base across context, population, mechanism, intervention, outcome and evaluation design, the study created a foundational structure through which cross-cutting pathways could be traced across

sectors while preserving the distinctions necessary to avoid false equivalence between studies that differ in what they examine, how they operationalise it, and the claims their designs can support.

2.2. Scoping study of the definition of community connectedness

What we did:

A scoping study examined how community connectedness is conceptualised and operationalised across relevant research traditions. Definitions, constructs and measures were mapped, leading to a bounded working definition: community connectedness refers to the patterned ways individuals and groups experience belonging, interaction, access and recognition within shared social, spatial and institutional contexts. To organise the evidence, four analytically distinct but connected domains were identified: individual or psychosocial, relational or interactional, place-based, and institutional or systems. Commonly used terms, including social capital, social cohesion, social integration and sense of community, were also considered in relation to these domains, drawing on established traditions in the field (McMillan & Chavis, 1986; Putnam, 2000).

What we learned:

The scoping study highlighted that similar labels can refer to different empirical content, while different labels may capture overlapping constructs. Evidence therefore needs to be classified according to what is actually measured rather than by terminology alone. For example, social integration may refer to interaction, access or perception depending on its operationalisation, while social cohesion may combine content across several domains. This reduces conceptual slippage and the risk of treating incompatible measures as equivalent. It also allows qualitative and quantitative evidence to be located within a common conceptual structure without assuming that different methods capture the same phenomenon. By separating terminology from the constructs and measures that sit beneath it, the scoping work created a standardised basis for locating otherwise heterogeneous evidence; this, in turn, sharpened the boundaries of review questions and inclusion decisions, while preserving an important distinction between describing what the evidence represents and making stronger claims about causality, effectiveness or policy relevance.

2.3. Desk review of existing What Works Centres

What we did:

Nine What Works Centres and related evidence organisations were reviewed to examine different approaches to evidence synthesis and use. The mapping included NICE, the Education Endowment Foundation, the College of Policing What Works Centre for Crime Reduction, Foundations, the What Works Centre for Local Economic Growth, the Centre for Ageing Better, the Centre for Homelessness Impact, the Wales Centre for Public Policy and the Youth Futures Foundation. The comparison focused on how these organisations defined their evidence domains, produced and communicated evidence, engaged decision-makers, and connected synthesis with practice and implementation. Wider work on the What Works movement provided additional context (Gough & Maidment, 2018; Davies et al., 2019).

What we learned:

The mapping identified several recurring features: clearly bounded domains, recognisable

evidence products, established relationships with intended users and, in several cases, implementation or improvement support. It also highlighted variation between evidence-heavy models centred on formal standards and evidence hierarchies, and more practice-embedded models emphasising co-production, tailored synthesis and applied guidance. The mapping highlighted three linked conditions shaping the usefulness of evidence synthesis: methodological credibility, relevance established when review questions are framed, and translation into clear, recognisable products that decision-makers can readily interpret and use (Gough & Maidment, 2018; Sanders & Breckon, 2023). For IU1, the comparison provided useful reference points for balancing methodological clarity with sufficient flexibility to accommodate heterogeneous evidence on complex social issues. Because the comparison showed that evidence organisations achieve credibility, relevance and practical influence through different combinations of methodological standards, domain boundaries, user relationships and forms of translation, its significance for IU1 lies in clarifying how an evidence-synthesis approach can retain sufficient analytical discipline to support defensible claims across a heterogeneous field while also shaping those claims, and the products through which they are communicated, around the conditions under which evidence can be understood and used in practice.

2.4. Stakeholder interviews with evidence reviewers

What we did:

We interviewed evidence reviewers (n=7) who have extensive experience in conducting systematic reviews and evidence syntheses in January-February 2026. More than half of these reviewers were based in What Works Centres (n=4), while the rest were either affiliated to institutes of higher education (n=1) or independent consultants (n=2).

The evidence reviewers were asked about their evidence review practices across four key areas: 1) the types of reviews they typically conduct; 2) the forms of evidence they include and challenges encountered, particularly around using grey literature and using qualitative evidence; 3) the frameworks and tools they use to assess evidence; and 4) the approaches they adopt to present evidence clearly and transparently.

What we learned:

1) Type of reviews conducted

The most common type of reviews conducted was **rapid reviews (n=6)**, often because policymakers and practitioners who commission these reviews require timely answers that fit within operational and policy cycles. Rapid reviews were often viewed as the next best alternative to **systematic reviews**, which remained highly valued and were undertaken by four stakeholders. Systematic reviews were generally viewed as the most rigorous method for synthesising evidence on intervention effectiveness and causal impact (n=4), but were perceived as expensive and time-consuming to do. Several stakeholders also emphasised the value of **scoping (n=2) and mapping (n=2) approaches to reviews**. These approaches were often described as important precursor activities that help identify the volume, nature, and distribution of evidence before investing in a full systematic review.

2) Forms of evidence included in reviews

Across the seven stakeholders interviewed, there was broad agreement that the type of evidence included should be driven by the purpose of the review rather than adherence to a single synthesis method. **Quantitative experimental and quasi-experimental studies**, including randomised controlled trials (RCTs) and quasi-experimental trials, were the most commonly

included forms of evidence (n=6). These designs were primarily used to answer questions about intervention effectiveness and "what works."

Nevertheless, many stakeholders argued that effectiveness evidence alone is insufficient for policy and practice decision-making. **Qualitative evidence** was valued for providing insights into implementation processes, mechanisms of change, contextual influences, user experiences, and feasibility considerations.

Stakeholders also defended the inclusion of grey literature within reviews because of its importance for understanding implementation, practice, and policy contexts. Nevertheless, several challenges were identified by stakeholders in regard to grey literature. The first challenge was the fragmented nature of grey literature searching, where reviewers must manually identify relevant organisations, search websites, and replicate database searches across multiple sources. This meant that grey literature searches can often be time and resource-intensive. The second challenge related to the quality and reporting of grey literature. Several stakeholders reported that grey literature frequently lacks sufficient methodological detail to permit robust quality appraisal.

3) Frameworks and tools used to assess evidence

There were two broad categories of frameworks used to assess evidence. One category focused on approaches that assess the suitability of types of evidence for inclusion, the second focused on approaches that evaluate the quality of individual primary studies.

For the first category, the most common approach to assess evidence was the **Maryland Scientific Methods Scale (SMS) (n=3)**, particularly amongst reviewers whose work focused on evaluating intervention effectiveness through experimental and quasi-experimental studies. Another stakeholder favoured the **Nesta Standards of Evidence (n=1)**, arguing that it is straightforward, intuitive, and useful for decision-makers.

For the second category, a diverse range of tools were mentioned for the appraisal of individual study quality across all seven stakeholders, including Cochrane Risk of Bias 2, the Mixed Methods Appraisal Tool (MMAT), the Critical Appraisal Skills Programme (CASP) checklists, and the Joanna Briggs Institute (JBI) critical appraisal tools. It has however been noted that each tool has limitations which must be taken into account when doing a review.

4) Approaches to present evidence

There was strong agreement that evidence should be communicated through a **portfolio of outputs tailored to different audiences**, rather than relying solely on traditional technical reports. Technical reports were generally still regarded as essential as they provide the methodological transparency and detail required for scrutiny and verification. The most commonly preferred reported output was toolkits (n=4), valued because they translate large bodies of evidence into practical decision-making tools that allow policymakers and practitioners to quickly identify effective interventions. Some stakeholders reported preferences for **evidence and gap maps (n=3)**, which were used to visualise the evidence landscape, identify research gaps, and inform future research commissioning, whilst others highlighted the importance of **plain English summaries (n=3)** to make evidence accessible to non-specialist users.

Dissemination and evidence mobilisation methods were viewed as equally important as the outputs themselves. The most common knowledge mobilisation approach was the use of **practitioner engagement events, workshops, and sessions (n=4)**. These activities were viewed as important mechanisms for translating research findings into practice, building relationships with evidence users, and creating opportunities for discussion and interpretation. Several

stakeholders also suggested that trust, engagement, and ongoing relationships with users are often more influential than the production of evidence outputs alone.

5) Use of AI within reviews

Even though stakeholders were not explicitly asked about their views on using artificial intelligence (AI) within evidence synthesis, it was brought up by more than half the reviewers (n=5). Attitudes towards AI were broadly positive but cautious. There was a strong consensus that AI has significant potential to improve the efficiency of evidence reviews, particularly for labour-intensive stages such as literature searching, screening, study classification, and data extraction. However, stakeholders consistently emphasised that AI should be viewed as a tool to support reviewers rather than replace human judgement, especially at the synthesis stage. Human oversight was regarded as particularly essential for interpreting evidence, understanding context, and ensuring that review findings are relevant to policy and practice needs.

2.5. Stakeholder interviews with evidence users

What we did:

We interviewed policymakers and practitioners (n=4) who have either used or were inclined to use systematic reviews to inform their decision-making processes in January-March 2026. These evidence users, chosen as part of a convenience sample, were affiliated to one of two institutions, a government department and a charity.

The evidence users were asked about how they engage with reviews and research syntheses in four key areas: 1) the types of research questions they hope to address with reviews; 2) the relative importance they place on different evidence types; 3) how they draw on evidence across different stages of the policy development processes; and 4) their preferences for evidence presentation formats and synthesis outputs.

What we learned:

1) Types of review questions

Across all four evidence users, reviews were perceived as being useful for answering applied, policy-relevant research questions that could directly inform real-world decision-making, resource allocation, and programme design within often tight policy timescales. The evidence users from the government department expressed significant interest in "what works" questions but also stated that reviews should examine the underlying causal drivers of problems, helping policymakers understand not just whether an intervention is effective but why outcomes occur and how communities can be engaged and mobilised effectively.

In contrast to the interviewed government department users who sought to use reviews to support immediate policy decisions, questions by the evidence users from the charity were often framed around understanding long-term community outcomes, strengthening local systems, and identifying the conditions that help communities flourish. Their research priorities were centred on identifying where public investment generates the greatest social value, which social and cultural institutions are most deserving of support, and what wider impacts community-based programmes have on outcomes such as health, employment, and education.

2) Use of different evidence types

Across the four stakeholders, there is little support for relying exclusively on a single research

design or type of evidence. Instead, they value evidence that can help answer different types of policy questions.

Interviewed evidence users from the government department valued causal and impact-focused evidence, including systematic reviews of experimental studies (e.g., RCTs and quasi-experimental designs), particularly when assessing whether policies and programmes are effective. However, they acknowledged that such methods were not always appropriate or feasible within complex policy environments and that other evidence types may be needed depending on the question being addressed. Qualitative studies and case studies were perceived as useful for understanding place-based contexts and local systems, but limitations in terms of generalisability of findings were raised, calling for careful interpretation and clear caveats where qualitative evidence is used.

In a similar vein, evidence users from the charity were proponents of a mixed-methods approach to evidence, drawing on qualitative, quantitative, and mixed-methods studies to understand the complex drivers of social cohesion. A central feature of their approach is also the use of qualitative evidence, which they draw upon to understand local contexts, community dynamics, lived experiences, and the processes through which cohesion is fostered or undermined. At the same time, they value quantitative evidence that can help identify patterns, outcomes, and trends across places, particularly as they develop tools such as the Social Cohesion Measurement Index.

3) How they draw on evidence at different stages of the policy process

Across all four stakeholders, evidence reviews were viewed as valuable at different stages of the policy cycle, although they were used in somewhat different ways depending on organisational context.

For evidence users within the government department, evidence can be drawn upon at multiple stages. At the problem-definition and agenda-setting stage, evidence from reviews could be used to understand the nature of policy challenges, identify evidence gaps, and inform strategic priorities. During the policy design and programme development stage, evidence could be used to inform decisions about interventions, implementation approaches, and community engagement strategies. Finally, evidence from reviews could be used to support spending reviews and business cases.

In contrast, evidence users from the charity explained that reviews are used primarily at the front end of the policy process to identify priorities, shape lines of enquiry, and inform the development of social cohesion strategies.

4) Preferences for evidence presentation formats and outputs

All four evidence users expressed a strong preference for evidence products that are accessible, concise, visually engaging, and tailored to the needs of different audiences. They consistently emphasised that evidence needs to be translated into formats that busy policymakers, local government officers, community leaders, neighbourhood boards, and practitioners can quickly understand and apply.

Evidence users from the government department found lengthy academic reports difficult for busy policymakers and practitioners to engage with. Some examples of preferred accessible presentation formats include slide decks, pre-recorded presentations/videos, posters, and short manuals. The evidence users stress the importance of communicating a clear "so what?" message, as well as layered or tiered evidence products, in which users are presented with headline findings first and can subsequently access outputs of progressively greater levels of

detail if required. They also emphasised the importance of active dissemination and engagement through mailing lists and targeted engagement events or activities.

Evidence users from the charity advocated for dashboard-style resources such as the Youth Endowment Fund Toolkit, which was perceived as an exemplar model because of its clarity, simplicity, and usability for non-specialist audiences. They suggested that a similar platform for social cohesion and community connectedness would be highly valuable. Their preferred outputs include evidence and gap maps, and concise, plain-English findings that can support decision-making across multiple levels of local government and community governance.

3. Overview of our approach

Section three of this framework document details C4 IU1's general approach to conducting systematic reviews and evidence syntheses for topics on community connectedness, in terms of research methodology, stakeholder and public engagement avenues, output types, and dissemination methods.

3.1. Guiding principles of our approach

- **Rigour:** We aim to conduct our reviews in a manner that is as comprehensive and thorough as possible. To this end, where relevant and appropriate, our reviews will consider evidence published in a wide range of sources, including academic databases, grey literature databases, and organisational websites. However, the extent of this comprehensiveness may be constrained by available time and resources. For example, due to limited resources, C4 IU1's reviews will only consider publications in English. This limitation and its implication(s) will be noted in the respective reports and outputs.
- **Proportionality:** Review methods will be selected according to the nature of the question, the characteristics of the available evidence, the intended use of the findings, and the time and resources available. C4 IU1 will seek the most rigorous approach that is appropriate and feasible within these parameters, rather than applying a single methodological model to all review questions. Any methodological restrictions, adaptations or efficiencies, including those made in rapid reviews, will be specified in advance wherever possible, documented transparently, and considered when interpreting the strength and scope of the resulting conclusions.
- **Transparency:** Our reviews will be reported in a transparent and reproducible manner, for example in accordance with the PRISMA reporting guidelines (Page et al., 2021) or other reporting guidance appropriate to the review design, where applicable. A protocol will also be registered before each review to detail the methodological approach.
- **Knowledge translation and usability:** The key purpose of our reviews is to support evidence-informed decision-making and contribute to the design and delivery of interventions that strengthen community connectedness, particularly within the UK context. Review outputs will be presented in formats that are user-friendly and accessible such that they address real-world needs and maximise usefulness and uptake.

3.2. Stages of reviews

3.2.1. Formulating the review question

While subject to change and variation, we foresee that C4 IU1 will predominantly seek to address three broad categories of review questions pertaining to community connectedness:

1) Review questions to map/define concepts or terminology

These review questions explore concepts, definitions, and evidence landscapes, seeking to understand how key ideas, terms, and models are defined and used within the literature, and to map the breadth and characteristics of the available evidence. Answering these questions will help to establish conceptual clarity and highlight gaps in the evidence base. Such questions are typically addressed using diverse types of evidence, incorporating quantitative, qualitative, and mixed methods evidence.

2) Review questions around effectiveness, often framed as “what works?” questions

These review questions examine whether interventions, programmes, policies, or approaches achieve their intended outcomes and, where possible, assess the magnitude and consistency of their effects. Such questions are typically addressed using quantitative evidence and may draw upon study designs that enable comparisons between interventions and outcomes.

In setting these review questions, the **PICO** (Population, Intervention, Comparison, Outcome) framework and its varied forms will be adopted as appropriate (Davies, 2011). These variations include **PICOT** (adds timeframe), **PICOC** (adds context/country), **PICOS** (adds study design). In cases where a review does not seek to evaluate a particular intervention but a condition or factors that individuals may come into contact with, **‘PECO’** (or another more appropriate variation) may be used instead whereby ‘I’ is replaced with ‘E’ for exposure (James et al., 2016).

3) Review questions around acceptability and contextual factors, often framed as questions on “why does x work, how does x work, for whom does x work, and under what circumstances?”

These review questions seek to explore experiences, perspectives, implementation processes, contextual influences, and mechanisms of change. Such questions are particularly important in complex social and community settings, where outcomes may be shaped by local contexts, stakeholder experiences, and wider social, economic, or institutional factors. Such questions are typically addressed using qualitative or mixed methods evidence and may draw upon study designs that provide rich and thick descriptions.

In these cases, variations of the **SPICE framework** (Setting, Perspective, Intervention or Phenomenon of Interest, Comparison, Evaluation) (Booth, 2006), the **SPIDER framework** (Sample, Phenomenon of Interest, Design, Evaluation, Research type) (Cooke et al., 2012) or the **PerSPecTIF framework** (Perspective, Setting, Phenomenon of interest/ Problem, Environment, Comparison (optional), Time/ Timing, Findings) (Flemming & Noyes, 2021) could be used to set the review question.

3.2.2. Registering protocols

In alignment with C4 IU1's guiding principles, a protocol for each review will be completed

and registered on suitable registration platforms (e.g., PROSPERO for systematic reviews or rapid systematic reviews, or Open Science Framework for scoping reviews, mapping reviews/evidence gap maps). Protocols reduce the risk of selective reporting, post hoc decision making, and bias in favour of positive findings. They also prevent unnecessary duplication of effort, support consistent and reproducible review processes, allowing any deviations from the initial plan to be clearly identified and justified (The PLoS Medicine Editors, 2011).

If changes to the protocol are deemed appropriate during the review, a section in the final report will be incorporated to shed light on what changes, if any, have been made from the final version of the protocol, and why.

3.2.3. Search and retrieval of studies

Each search conducted aims to be both sensitive (i.e., retrieve a high proportion of relevant studies) and specific (i.e., retrieve a low proportion of irrelevant studies) (Petticrew & Roberts, 2006).

Search sources:

Search sources will be selected according to the review question, relevant disciplines, study designs and types of evidence sought. Reviews will normally search multiple bibliographic databases, combining appropriate multidisciplinary and subject-specific sources where required. The selection of databases and other sources will be justified and documented in the review protocol, and may include:

- SCOPUS
- Web of Science (which includes the Social Science Citation Index (SSCI))
- IBSS (International Bibliography of the Social Sciences)
- ProQuest (which includes the Applied Social Sciences Index and Abstracts (ASSIA))

To ensure that the searches are specific and fit-for-purpose, additional subject databases (including but not limited to the ones listed below) may also be searched accordingly, depending on the topical focus and research question(s) (Papaioannou et al., 2010; Pickup et al., 2018).

- Economics: EconLit
- Education: British Education Index, ERIC [Education Resources Information Centre Database]
- Information Technology: IEEE Xplore
- Medicine: PubMed/MEDLINE, Embase
- Nursing: CINAHL [Cumulative Index to Nursing and Allied Health Literature]
- Psychology: PsycINFO

In alignment with our guiding principles, the need for grey literature searching will be considered for each review according to the review question and evidence sought. Where grey literature is likely to provide substantively important evidence, appropriate sources will be identified and searched using a transparent and reproducible strategy as far as practicable. This will use

similar (or, where possible, the same) search terms as those used for the academic databases. Targeted organisation or government websites (e.g., Ministry of Housing, Communities and Local Government) and search engines (e.g., Google) will be hand searched for relevant reports. Doctoral dissertations may also be searched on Dissertations & Theses Global (ProQuest).

Search strategy:

Search concepts underpinning the research question(s) will be identified and populated with the relevant search terms. These search concepts will be systematically derived from the structured framework(s) used to establish the research question (see sub-section 3.2.1).

The Boolean operators OR and AND will be used to connect search terms within a concept and between search concepts respectively. To maximise the sensitivity of the search, where appropriate, truncations and wildcards (* or ?) will be used to represent variations of different terms (e.g., plurality, American vs British spelling), while proximity operators such as Near (Nn) and Within (Wn) operators may also be used to enhance search strings (McGill University Library, 2026).

3.2.4. Study screening and selection

How criteria are developed: The inclusion and exclusion criteria will be determined based on the structured frameworks used to establish the research question (see sub-section 3.2.1).

Screening stages: Each review will undergo two successive screening stages, 1) Screening on title and abstract; 2) Screening on full text. Only studies that have been included at the title and abstract stage will be retrieved for full text screening.

Piloting the screening tool: At the beginning of each screening stage, a purposive sample of records will be independently screened by members of the review team to test the interpretation and application of the eligibility criteria. The size of the pilot will be proportionate to the size, complexity and heterogeneity of the evidence base. Disagreements or ambiguities will be discussed and the criteria refined where necessary before substantive screening proceeds.

Dual screening and reconciliation: Where time and resources allow, studies within the review will be independently screened by two reviewers at both screening stages. This improves reliability in the application of inclusion/exclusion criteria by reducing selection bias as well as ensuring transparency and reproducibility. Disagreements will be discussed and reconciled between the reviewers, and discrepancies that cannot be resolved will be directed to a third reviewer.

Where full independent dual screening is not feasible, the review protocol will specify a proportionate alternative approach. This may include independent duplicate screening of a sample of records following piloting, supplemented by ongoing quality assurance and referral of uncertain cases to a second reviewer. The approach adopted, and its implications for the review, will be reported transparently.

3.2.5. Data extraction

For included studies, data will be extracted using a structured extraction form developed by the review team (Taylor et al., 2021). The exact data extraction form for each review will be highly dependent on the review question(s) and objective(s), but some possible data extraction fields are listed below.

Possible data extraction fields for all reviews: Regardless of review type, question(s) or topic, the categories of characteristics listed in Table 1 may be extracted for all reviews (James et al., 2016).

Table 1

Category	Examples of Data Items
General	• Authors • Year of publication • Research design
Population / Sample	• Age/age group • Gender/sex • Ethnicity • Socio-demographic group • Profession
Context	• Country/geographical context • Cultural context

Possible data extraction fields for reviews on intervention effectiveness or “what works” reviews: For reviews on interventions or “what works”, the categories of characteristics listed in Table 2 may be extracted as appropriate (Brown et al., 2003; Elamin et al., 2009).

Table 2

Category	Examples of Data Items
Intervention / Exposure	• Type(s) of intervention/exposure • Components/approaches of the intervention/exposure • Length and duration of the intervention/exposure • How the intervention was administered
Comparator	• Type(s) of comparator (e.g., treatment as usual, placebo, alternative intervention) • Components/approach of the comparator • Length and duration of the comparator • How the comparator was administered (if relevant)
Study Design	• Sample size • Blinding procedures • Randomisation process/ assignment to groups
Outcomes	• Outcome type • Mean values for the outcomes of the experimental and control groups • Standard deviations (SDs) or standard errors of the mean (SEMs) • Sample sizes (N) per group • Measurement time points

Possible data extraction fields for reviews on acceptability and contextual factors: For reviews on acceptability and contextual factors, the categories of characteristics listed in Table 3 may be extracted as appropriate (Noyes et al., 2024).

Table 3

Category	Examples of Data Items
Sample	Relevant background characteristics (e.g., lived experience)
Phenomenon	Views and experiences related to the review question and phenomenon of focus
Study Design and Methods	- Sample recruitment methods - Data collection methods - Theoretical models used to interpret or contextualize the findings

3.2.6. Quality appraisal and confidence assessment

Possible quality appraisal tools for use:

The most updated version of the Cochrane Risk of Bias tool will be used to critically appraise experimental studies such as randomised controlled trials. The tool has been recognised as a best-practice instrument for assessing risk of bias (Bilotta et al., 2014). Additionally, according to Wells and Littell (2009), the tool is grounded in empirical evidence and is conceptually sound.

For other study designs, the respective checklists designed by the Joanna Briggs Institute (JBI) or the Critical Appraisal Skills Programme (CASP) will be used as a base for quality appraisal.

For mixed methods reviews incorporating diverse types of evidence, the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018) or the Weight of Evidence tool (Gough, 2007) may also be used.

How appraisal informs synthesis and interpretation:

Critical appraisal will be used to inform the synthesis and interpretation of findings rather than serving only as an assessment of individual studies and rule on their inclusion. Studies will not normally be excluded solely on the basis of appraisal unless exclusion criteria established in the protocol require this. Instead, appraisal findings will be considered when determining the weight placed on particular evidence, exploring variation between findings, undertaking sensitivity analyses where appropriate, and judging the strength and limitations of the overall evidence.

The approach used to move from study-level appraisal to conclusions about the strength or confidence of the evidence will be appropriate to the review design and specified in the protocol. C4 IU1 will avoid presenting apparently precise conclusions where weaknesses, inconsistency or gaps in the underlying evidence do not warrant them.

Assessing confidence in the body of evidence:

Where appropriate, C4 IU1 will make an explicit assessment of the confidence or certainty

that can reasonably be placed in the body of evidence supporting key findings. Judgements may consider factors including methodological limitations, relevance to the review question and context, coherence or consistency of findings, precision and the adequacy and richness of the available evidence. Where established approaches to assessing confidence or certainty are appropriate to the review design, these may be used (e.g., GRADE framework).

Assessments of confidence will be communicated to enable users to distinguish between well-supported conclusions, findings that should be interpreted cautiously, and areas where the available evidence is insufficient to support a conclusion.

3.2.7. Synthesis

Choosing the synthesis method: The synthesis methods will likely be chosen based on the purpose of the review, as well as data extracted in response to the review question.

As a general rule, in accordance with Gough and Thomas (2016), for questions around effectiveness or “what works?”, the synthesis approach will be more **aggregative**, using statistical or structured methods to combine consistent findings to test hypotheses and estimate effect sizes. In contrast, questions around acceptability and contextual factors, often framed as questions on “why does x work, how does x work, for whom does x work, and under what circumstances?”, require more **configurative** synthesis methods, where diverse findings are organised and interpreted to answer questions. Where appropriate, we will also use **mixed methods syntheses**, combining both aggregative syntheses (which address empirical effectiveness) and configurative syntheses (which explore experiences, mechanisms, or contextual factors).

Types of possible synthesis methods:

More aggregative synthesis methods:

- **Meta-analysis:** A meta-analysis produces a summary estimate with a corresponding confidence interval using statistical pooling methods (Vetter, 2019). Where studies are sufficiently comparable in terms of population, intervention or exposure, comparator, outcomes and study design, statistical meta-analysis may be undertaken to estimate a summary effect with an appropriate measure of uncertainty. Decisions about whether studies can meaningfully be pooled will be based primarily on substantive and methodological considerations rather than on a single statistical threshold. Statistical heterogeneity will be examined using appropriate measures, which may include the I² statistic. Wherein, if the I² value meets a determined threshold, such as 40% or below, the pooled effect size will be presented as a finding of the review (Deeks et al., 2024). If the I² value exceeds 40%, alternative methods may be considered, such as narratively summarising the findings.

More configurative synthesis methods:

- **Thematic summaries:** Thematic summaries organise studies into thematic groups based on a conceptual framework established at the outset of the review. The aim is to provide a structured summary of findings under key themes relevant to the policy or practice question being addressed (Snilstveit et al., 2012).
- **Thematic synthesis:** We will adopt the steps outlined by Thomas and Harden’s (2008) seminal work on thematic synthesis. The steps taken will be as such: 1) line by line inductive coding to capture the meaning and content of each sentence of data extracted; 2) these initial codes will be examined and organised into descriptive themes that remain closely aligned with the primary studies’ reported findings; 3) descriptive themes will be analysed

and interpreted as analytical themes in relation to the review question, generating new insights that extend beyond the content of the original studies.

- **Framework synthesis:** As per Carroll et al. (2011), a pre-defined framework is first chosen for the synthesis. The data extracted are deductively coded into this framework. Any data that do not fit these initial concepts within the framework are then analysed inductively using thematic analysis to generate new themes, which are added to the framework. Together, the a priori and newly derived themes form an expanded conceptual structure that is further refined by examining how themes relate to one another, enabling reviewers to build a revised conceptual model.
- **Realist synthesis:** Data extracted will be interpreted as part of Context–Mechanism–Outcome (CMO) configurations, which will help to inform or develop existing programme theory (Duddy & Wong, 2023). Context refers to the conditions or circumstances that influence whether and how a mechanism will be triggered. Mechanisms are the underlying generative processes—the reasoning, reactions, or changes in participants—that lead to outcomes. Outcomes are the observable effects (intended or unintended) that result when mechanisms are triggered within specific contexts.

Both types/Mixed methods synthesis:

Where appropriate, we will also use mixed methods designs, where for example, an aggregative synthesis addresses empirical effectiveness, and a configurative synthesis explores experiences, mechanisms, or context.

A mixed methods synthesis will be conducted using the relevant convergent or sequential synthesis design specified in Hong et al. (2017). In convergent synthesis designs, qualitative and quantitative data are collected and analysed within the same phase of the review, with three variations: 1) data based convergent synthesis, where all data extracted is analysed together using a single synthesis method (i.e., either qualitative or quantitative), requiring data transformation to produce an integrated output; 2) results based convergent synthesis, where all data extracted are analysed separately using different synthesis methods, and then combined in a further integrative step as a third stage synthesis; and 3) parallel results convergent synthesis, where separate qualitative and quantitative syntheses remain independent and are only brought together during interpretation in the discussion section. In contrast, sequential synthesis designs involve two distinct phases, with the findings from the first synthesis—whether qualitative or quantitative—informing the focus, analysis, or interpretation of the second, enabling reviewers to build progressively from one evidence stream to the next.

3.3. Selecting review approach types

The choice of review approach will be determined primarily by the review question and intended use of the evidence rather than by a predetermined preference for a particular review type. In agreeing the approach, C4 IU1 will consider:

- The nature of the question being asked;
- The type, volume, maturity and heterogeneity of the likely evidence base;
- Whether the principal purpose is to map evidence, estimate effects, explain mechanisms or contexts, understand experiences, or combine these purposes;
- The degree of comprehensiveness and certainty required;

- How and when the findings will be used;
- Opportunities for stakeholder and community involvement;
- The time and resources available.

These considerations will inform the choice of systematic, scoping or mapping review and whether the review is undertaken in rapid mode. The chosen approach, and the rationale for it, will be documented in the review protocol.

3.3.1. Clarifying review approach types

Considerable confusion exists regarding the terminology and distinctions between different review approaches. For clarity and consistency, C4 IU1 adopts the following categorical definitions of the review types and what each type entails.

- **Systematic reviews (Nussbaumer-Streit et al., 2025):**

Purpose: To provide a comprehensive and reliable summary of the available evidence relating to a clearly specified review question by synthesising findings across all eligible studies.

Overall Methodology: Systematic reviews systematically identify, appraise and synthesise all available evidence, applying rigorous and reproducible methods to minimise bias. Systematic reviews can take the form of quantitative synthesis (e.g., via a meta-analysis), qualitative evidence synthesis or mixed methods synthesis.

Stages: All stages listed in section 3.2.

- **Scoping reviews (Campbell et al., 2023):**

Purpose: To clarify key concepts/definitions in the literature and identify key characteristics or factors related to a concept, including those related to methodological research.

Overall Methodology: Scoping reviews systematically identify and map the breadth of evidence available on a particular topic, field, concept, or issue, often irrespective of source (i.e., primary research, reviews, non-empirical evidence) within or across particular contexts. They are generally more exploratory and flexible in their approach, using inductive or iterative approaches to synthesis.

Stages: All stages in section 3.2; critical appraisal is not routinely required for scoping reviews but may be undertaken where it contributes directly to the purpose of the review.

- **Mapping reviews (Campbell et al., 2023):**

Purpose: To describe and catalogue/categorise evidence within a broad topic area, providing a broad overview of what evidence exists.

Method: Mapping reviews employ predefined coding frameworks and extract relatively high-level descriptive information from studies. This means that reviewers systematically categorise studies according to characteristics such as topic area, population, intervention type, setting, or outcome. They are generally more structured and descriptive in their approach.

Stages: All stages in section 3.2; critical appraisal will not normally be undertaken unless required by the purpose of the map.

- **Rapid reviews (Nussbaumer-Streit et al., 2025):**

Unlike the above review types, rather than representing a distinct type of review, a rapid review is best understood as a review mode that can be applied to any of the aforementioned review types (i.e., rapid systematic review, rapid scoping review etc.).

Purpose: To provide decision-makers with the necessary evidence within constrained timeframes or resources, thereby informing urgent policy, practice, or service decisions.

Method: Rapid reviews accelerate the evidence synthesis process through the use of carefully selected methodological shortcuts while maintaining a systematic and transparent approach.

Stages: Depends on the review type.

3.4. Engagement and co-production

Stakeholder and community involvement will be considered throughout the evidence review and synthesis process rather than solely at the point of dissemination. The nature and extent of involvement will vary according to the review question, intended users, timescale and resources available.

This aspect of the framework is being developed further alongside C4's wider approach to collaboration and community involvement.

3.5. Use of artificial intelligence (AI) within reviews

Artificial intelligence (AI) tools may be used to support appropriate stages of evidence reviews and evidence synthesis activities. The use of AI will be considered on a review-specific basis and may assist with activities such as developing search terms, refining search strategies, screening records, identifying relevant literature, extracting information from studies, summarising evidence, and drafting sections of review outputs. However, AI will not replace the judgement, expertise, or decision-making responsibilities of the review team. All AI-assisted activities will be subject to appropriate human oversight, verification, and quality assurance procedures to ensure the accuracy, validity, and integrity of review findings. The use of AI within reviews will be documented transparently, including where and how it has been applied, in order to support accountability and reproducibility. Furthermore, all use of AI will comply with relevant ethical, institutional, and data protection requirements, including policies relating to confidentiality, intellectual property, responsible research conduct, and the secure handling of data.

3.6. Report structure and output types

Each review will be documented in a technical report detailing methods, findings, limitations, and evidence quality, ensuring transparency and credibility for researchers, analysts, and technical users.

Additionally, C4 IU1 aims to produce a suite of complementary evidence products to cater to the target audience. This may include one or more of the following:

- Plain English summary;
- Policy brief;
- Slide deck;
- Infographic(s) and visual content;
- Practitioner guide and/or implementation checklist.

The selection and design of outputs will be informed by the intended users of each review. C4 IU1 will adopt a layered approach to dissemination, including through engagement activities such as webinars, workshops, and learning events. The intention is to enable users to access headline findings and implications quickly while retaining access to progressively greater methodological and evidential detail. Outputs will distinguish clearly between what the evidence shows, the level of confidence that can reasonably be placed in the findings, important evidence gaps, and the implications that may follow for policy, practice and future research.

References

- Bilotta, G. S., Milner, A. M., & Boyd, I. L. (2014). Quality assessment tools for evidence from environmental science. *Environmental Evidence*, 3, Article 14. <https://doi.org/10.1186/2047-2382-3-14>
- Booth, A. (2006). Clear and present questions: Formulating questions for evidence based practice. *Library Hi Tech*, 24(3), 355–368. <https://doi.org/10.1108/07378830610692127>
- Brown, S. A., Upchurch, S. L., & Acton, G. J. (2003). A framework for developing a coding scheme for meta-analysis. *Western Journal of Nursing Research*, 25(2), 205–222. <https://doi.org/10.1177/0193945902250038>
- Campbell, F., Tricco, A. C., Munn, Z., Pollock, D., Saran, A., Sutton, A., White, H., & Khalil, H. (2023). Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): The same but different—the “Big Picture” review family. *Systematic Reviews*, 12, Article 45. <https://doi.org/10.1186/s13643-023-02178-5>
- Carroll, C., Booth, A., & Cooper, K. (2011). A worked example of “best fit” framework synthesis: A systematic review of views concerning the taking of some potential chemopreventive agents. *BMC Medical Research Methodology*, 11, Article 29. <https://doi.org/10.1186/1471-2288-11-29>
- Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: The SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research*, 22(10), 1435–1443. <https://doi.org/10.1177/1049732312452938>
- Davies, K. S. (2011). Formulating the evidence based practice question: A review of the frameworks. *Evidence Based Library and Information Practice*, 6(2), 75–80. <https://doi.org/10.18438/B8WS5N>
- Davies, H., Nutley, S., Boaz, A., & Fraser, A. (Eds.). (2019). *What works now? Evidence-informed policy and practice*. Policy Press.
- Deeks, J. J., Higgins, J. P. T., Altman, D. G., McKenzie, J. E., & Veroniki, A. A. (2024). Chapter 10: Analysing data and undertaking meta-analyses. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, et al. (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.5). Cochrane. <https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-10>
- Duddy, C., & Wong, G. (2023). Grand rounds in methodology: When are realist reviews useful, and what does a “good” realist review look like? *BMJ Quality & Safety*, 32(3), 173–180. <https://doi.org/10.1136/bmjqs-2022-015236>
- Elamin, M. B., Flynn, D. N., Bassler, D., Briel, M., Alonso-Coello, P., Karanickolas, P. J., Guyatt, G. H., Malaga, G., Furukawa, T. A., Kunz, R., Schünemann, H., Murad, M. H., Barbui, C., Cipriani, A., & Montori, V. M. (2009). Choice of data extraction tools for systematic reviews depends on resources and review complexity. *Journal of Clinical Epidemiology*, 62(5), 506–510. <https://doi.org/10.1016/j.jclinepi.2008.10.016>
- Flemming, K., & Noyes, J. (2021). Qualitative evidence synthesis: Where are we at? *International Journal of Qualitative Methods*, 20, Article 1609406921993276. <https://doi.org/10.1177/1609406921993276>

- Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213–228. <https://doi.org/10.1080/02671520701296189>
- Gough, D., & Maidment, C. (2018). UK What Works Centres: Aims, methods and contexts. UCL Institute of Education. <https://eppi.ioe.ac.uk/cms/publications/primary-research/uk-what-works-centres-aims-methods-and-contexts>
- Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic Reviews*, 6, Article 61. <https://doi.org/10.1186/s13643-017-0454-2>
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O’Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). Mixed methods appraisal tool (MMAT), version 2018. Canadian Intellectual Property Office. Microsoft Word - MMAT_2018_criteria-manual_2018-08-08.docx
- James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence*, 5, Article 7. <https://doi.org/10.1186/s13750-016-0059-6>
- McGill University Library. (2026, July 27). Systematic Reviews, Scoping Reviews, and Other Knowledge Syntheses. <https://libraryguides.mcgill.ca/c.php?g=695878&p=5422760>
- McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1), 6–23. [https://doi.org/10.1002/1520-6629\(198601\)14:1%3C6::AID-JCOP2290140103%3E3.0.CO;2-I](https://doi.org/10.1002/1520-6629(198601)14:1%3C6::AID-JCOP2290140103%3E3.0.CO;2-I)
- Noyes, J., Booth, A., Cargo, M., Flemming, K., Harden, A., Harris, J., Garside, R., Hannes, K., Pantoja, T., & Thomas, J. (2024). Chapter 21: Qualitative evidence (last updated October 2019). In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, et al. (Eds.), *Cochrane handbook for systematic reviews of interventions* (Version 6.5). Cochrane. <https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-21>
- Nussbaumer-Streit, B., Booth, A., Garritty, C., Hamel, C., Munn, Z., Tricco, A. C., & Pollock, D. (2025). Overview of evidence synthesis types and modes. *Journal of Clinical Epidemiology*, 187, Article 111970. <https://doi.org/10.1016/j.jclinepi.2025.111970>
- 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., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Papaioannou, D., Sutton, A., Carroll, C., Booth, A., & Wong, R. (2010). Literature searching for social science systematic reviews: Consideration of a range of search techniques. *Health Information & Libraries Journal*, 27(2), 114–122. <https://doi.org/10.1111/j.1471-1842.2009.00863.x>
- Petticrew, M., & Roberts, H. (2006). How to find the studies: The literature search. In *Systematic reviews in the social sciences: A practical guide* (pp. 79–124). Blackwell Publishing. <https://doi.org/10.1002/9780470754887.ch4>

- Pickup, D. I., Bernard, R. M., Borokhovski, E., Wade, A. C., & Tamim, R. M. (2018). Systematically searching empirical literature in the social sciences: Results from two meta-analyses within the domain of education. *Rossiiskii Psikhologicheskii Zhurnal (Russian Psychological Journal)*, 15(4), 245–265. <https://doi.org/10.21702/rpj.2018.4.1>
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon & Schuster.
- Sanders, M., & Breckon, J. (2023). *The What Works Centres: Lessons and insights from an evidence movement*. Policy Press.
- Snilstveit, B., Oliver, S., & Vojtkova, M. (2012). Narrative approaches to systematic review and synthesis of evidence for international development policy and practice. *Journal of Development Effectiveness*, 4(3), 409–429. <https://doi.org/10.1080/19439342.2012.710641>
- Taylor, K. S., Mahtani, K. R., & Aronson, J. K. (2021). Summarising good practice guidelines for data extraction for systematic reviews and meta-analysis. *BMJ Evidence-Based Medicine*, 26(3), 88–90. <https://doi.org/10.1136/bmjebm-2020-111651>
- The PLoS Medicine Editors. (2011). Best practice in systematic reviews: The importance of protocols and registration. *PLoS Medicine*, 8(2), e1001009. <https://doi.org/10.1371/journal.pmed.1001009>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, Article 45. <https://doi.org/10.1186/1471-2288-8-45>
- Vetter, T. R. (2019). Systematic review and meta-analysis: Sometimes bigger is indeed better. *Anesthesia & Analgesia*, 128(3), 575–583. <https://doi.org/10.1213/ANE.00000000000004014>
- Wells, K., & Littell, J. H. (2009). Study quality assessment in systematic reviews of research on intervention effects. *Research on Social Work Practice*, 19(1), 52–62. <https://doi.org/10.1177/104973150831727>