EDUCATION
.ORG

SYNTHESES & REPORTS

PUBLISHED ON
MAY 29, 2025

Why and how to use a more inclusive evidence base to inform policy and practice

Guidance for widening the evidence base in education decision-making

International Working Group (IWG)

ASK THE EVIDENCE

GUIDANCE for widening the evidence base in education decision-making

Framing the analytical question

- Why framing a good analytical question is important
- Types of analytical questions
- What a good analytical question looks like

Identifying and accessing sources

- Purpose of developing a systematic search plan
- Process of developing and implementing a systematic search plan

Appraising sources

- Purpose of appraising sources
- Introduction to the Appraisal Criteria
- Process for using the Appraisal Tool
- Considerations for adapting the Appraisal Tool

Introduction

Why does education decision-making need a broader evidence base?

Despite consensus on the scale of the learning crisis, the global education sector continues to struggle to identify and scale context-sensitive solutions. Education decision-makers are actively seeking better evidence to inform these efforts, yet the conventional evidence base—dominated by peer-reviewed, academic research— often fails to meet their need for practical, timely, and context-relevant guidance.

The omission of large volumes of non-academic work of great value – so-called 'grey literature' – severely limits the available evidence base. As described in the 2021 Education.org White Paper: Building an Education Knowledge Bridge, this also limits the perspectives represented in evidence, since much of this grey literature is produced outside of traditional academic publishing channels, including by NGOs, government agencies, civil society organisations, and practitioners—many of whom are embedded in the very contexts policies aim to serve.

What's wrong with how we define and access evidence today?

Current definitions of 'good evidence' overly privilege academic research, often published in English-language, international journals and behind paywalls. This narrows access to insights as well as the scope of perspectives represented—particularly those from the Global South—and excludes the experiential insights of educators and local communities. Many practitioners and policymakers report difficulty accessing research or trusting guidance that lacks contextual relevance or practical orientation.

Grey literature is increasingly recognised for its potential to fill these gaps, yet many systems still exclude it from formal evidence reviews. According to an Education.org landscaping exercise, 'Why Is So Much Useful Education Evidence Still Overlooked?' to understand the use of grey literature, only 9 of 26 international initiatives reviewed included it in their evidence definitions. Even then, most imposed academic-style quality standards that many of these sources were never designed to meet.

In Education.org's synthesis providing insights on improving successful transitions from Accelerated Education to formal schooling, nearly 80% of the useful sources were unpublished, non-academic literature, which resulted in a doubling of the number of countries included.

The health sector has recognised the value of grey literature for its documentation of practitioner experiences and local context, which helps indicate how and why things work (Kothari et al. 2012; Lewin et al. 2024).

> " Expanding the range of evidence sources to include grey literature adds tremendous value by enriching contextual relevance, valuing practice-derived lessons, and expanding the perspectives to include the work of researchers and practitioners in the Global South. "

Education.org

How do we define 'grey literature'?

The most widely accepted definition of 'grey literature' is the 2010 Prague definition:

> " Grey literature stands for manifold document types produced on all levels of government, academics, business and industry in print and electronic formats that are protected by intellectual property rights, of sufficient quality to be collected and preserved by libraries and institutional repositories, but not controlled by commercial publishers; i.e. where publishing is not the primary activity of the producing body (Schöpfel 2010). "

This Guidance include a slightly wider scope to include products from think tanks, political parties, funding agencies, international non-governmental organisations (INGOs), civil society organisations, and schools and other educational institutions, regardless of their governance. These sources may be research. But they may also not be research; they may be descriptive, analytic, or first-hand accounts providing contextual information relevant to decision-makers.

How does this Guidance aim to address the problem?

Education.org and its International Working Group (IWG)—a diverse network of 27 organisations across multiple continents—developed this Guidance to build on existing appraisal guidance and to develop a new, coherent, and intellectually rigorous system for making a wider range of evidence easier to identify, access, and use in educational decision-making (see Appendix A for a list of members).

Through a year-long, iterative process, the IWG collectively developed the Guidance and held consultations in a variety of forums. The draft Guidance was tested in Kenya and Sierra Leone, then revised by Education.org to support analysts, researchers, and evidence users in three key areas:

1. **Framing analytical questions:** Guidance on crafting relevant, policy-facing questions that reflect the concerns and priorities of decision-makers

2. **Identifying and accessing evidence**: A systematic approach to search for both academic and grey literature, emphasising inclusion of local and underrepresented perspectives.

3. **Appraising evidence for use:** A novel appraisal tool designed to assess relevance, inclusivity, credibility, and methodological rigour across diverse source types.

This Guidance also builds on the LEARRN partnership's Evidence Classification Framework, offering more technical tools for those conducting research and analysis.

Who is this Guidance for?

While researchers and evidence synthesisers are its primary audience, the Guidance is also designed to inform actors within the broader education sector.

Users and uses of this Guidance

Research and evaluation

Researchers

While academics and other education researchers often speak of or use grey literature, this Guidance makes it easier to find them and provides a stronger rationale for including them as credible sources, including in publications. The application of this Guidance by a range of researchers can help redefine who are 'experts' and amplify the work of locally generated knowledge producers, including early career researchers.

Evidence synthesisers

This Guidance can assist evidence synthesisers to use a more inclusive set of sources and potentially increase the value of syntheses in addressing decision-maker questions. The appraisal tool can be integrated in their protocol.

Evaluators

This Guidance facilitates access to local experiences and provides a rigorous means for appraising it with attention to potential bias.

Non-research

Government actors (analysts, policy advisors, etc.)

This Guidance provides a time-saving road map to access contextually relevant sources, a pathway that can inform longer-term mechanisms for strengthening the evidence – policy/planning nexus. Those with less experience in some research designs, may find appraising some types of research outside their expertise. However, their analyses will be enhanced by applying the criteria of relevance, credibility and inclusivity and considering the limitations and bias of sources for all evidence.

Practitioners

Because practitioners are also knowledge producers, familiarity with this Guidance can influence their writing to better serve education decision-makers and suggest ways to expand access to their knowledge products. The appraisal criteria of relevance, credibility and inclusivity provide key principles for their work, even if they do not use the appraisal tool.

Consultants

This Guidance provides time-pressed consultants with a jump-start in their search for local experience and expertise and a system for mitigating bias.

Funders

Funders have tremendous influence over the research, analyses and programme documentation that they fund. The Guidance can help bring funders' own knowledge products to more users and promote the work and voices of those they seek to serve. They can ask researchers to include sources beyond published research. They can also make the reports they fund (evaluations, research, and programme reports) more accessible so that the richness of the work is built upon. Funders should push knowledge producers to make their work accessible through OpenAlex: https://openalex.org.

When should this guidance be used?

This Guidance is most useful in the early stage of designing research reviews or conducting a secondary analysis. It guides the user in framing questions from an education decision-maker's perspective. It helps point to evidence sources that may not have considered. This contributes to understanding what is known and helps determine if more data needs to be collected. Once evidence has been attained, the Guidance then helps in assessing the quality and relevance of those individual sources. This Guidance does not assist in the analysis or in the design of a new piece of primary research.

What impact do we hope to achieve?

By promoting inclusive, systematic methods for using both academic and non-academic sources, this Guidance aims to:

- Elevate diverse, often-marginalised voices in evidence-informed education policy;
- Increase the use and visibility of locally generated insights;
- Promote transparency, trust, and relevance in policy-relevant research.

Ultimately, the goal is to strengthen the evidence-policy-practice bridge, creating a world where education outcomes are improved due to better decision-making, which is informed by the most relevant and useful evidence available.

For further information, editable tools, and training, contact: info@education.org

Part 1. Framing the analytical question

Why framing a good analytical question is important

In the context of evidence-based policymaking, an analytical question is a question that asks you to examine data, compare information or results, and explain the reasons behind certain outcomes or decisions. Instead of just asking for facts, it challenges you to explore *why* something happened, *how* different factors are linked, or *what* solutions might be most effective. Framing an analytical question is a critical first step in the process of identifying, accessing, and appraising evidence. A well-crafted question is essential because:

- **It guides subsequent steps,** shaping what sources of information are most relevant, how to analyse them, and what conclusions can be drawn.

- **It helps you stay focused,** deciding which sources and what type of information to look for, ultimately ensuring your findings are useful and applicable in a given context.

- **It ensures a solution-oriented approach,** collecting evidence on a specific question, urgent challenge, or pressing policy topic to better inform education decision-makers.

This is especially important when trying to inform decision-making processes. Without a strong analytical question, research or analysis can become too broad, unfocused, or irrelevant, making it difficult to generate clear, practical conclusions or actionable ways forward for decision-makers.

What are the different types of analytical questions

An analytical question informs how information is scrutinised in order to generate insights for policy or practice. For example, a question can be descriptive, explanatory, exploratory, comparative, or evaluative. The table below provides a brief explanation and example questions for each of these types of analytical questions.

TYPE OF ANALYTICAL QUESTION	EXAMPLES
Descriptive questions aim to provide details about a phenomenon, event, or trend. They usually attempt to answer 'what' is happening, in order to document factors or characteristics without analysing causes and effects.	• What are the documented barriers to teacher retention in rural primary schools in Sub-Saharan Africa? • How do dropout rates differ amongst students, based on student gender, disability, and household location in Latin America?
Explanatory (or causal) questions investigate the reasons behind a phenomenon and often involve cause-and-effect relationships. These questions help determine whether one factor influences another.	• What impact does school feeding have on school attendance rates in low-income countries? • How did the COVID-19 pandemic impact girls' access school in lower- and middle-income countries?
Exploratory questions help uncover patterns, generate new ideas, especially on a topic that is not well known. The goal is to gain a deeper understanding.	• What are teachers' perceptions of the new competency-based curriculum in country X? • How do students and their families perceive non-formal education in the Middle East?
Comparative questions focus on differences and similarities between groups, settings, or time periods. They help assess variations and trends.	• What teaching strategies are used in Kenyan schools and non-formal education models? • How do government-led and NGO-led school feeding programmes differ in their impact on student attendance in Colombia?
Evaluative questions assess the effectiveness of programmes, policies, or interventions. The purpose is to measure outcomes and inform improvements.	• What evidence exists on the effectiveness of community-based early childhood education programmes in improving literacy outcomes in lower- and middle-income countries? • How effective are government-led mentorship programmes in improving school completion rates for girls in Pacific Island Countries?

Analytical questions do not always fit neatly into just one of the types listed above. Many questions overlap, combining elements of multiple types of questions to examine a topic in a deeper, more nuanced way. Below are a few examples to illustrate this:

Example question #1: How do differences in school infrastructure across urban and rural areas influence primary student attendance in sub-Saharan Africa?

- The question is descriptive because it documents what school infrastructure looks like
- The question is comparative, because it examines differences between infrastructure in rural and urban areas
- The question is also explanatory, because it seeks to identify causal relationship between school infrastructure and its impact on student attendance.

Example question #2: What factors contribute to the success of non-formal community-based education programmes in fragile and conflict-affected countries, and how do they impact student learning outcomes?

- The question is exploratory because it seeks to identify success factors.
- The question is explanatory because it connects the cause-and-effect relationship between the success factors and student learning.
- The question is evaluative because it assesses the impacts of community-based education models.

Note that you may want to craft multiple analytical questions to capture the different purposes of your study. For example, you may have a main question, with various sub-questions, such as:

What factors contribute to the success of community-based education programmes?

- How do these factors impact student learning in literacy and mathematics?
- How do these factors impact student social-emotional learning?
- What barriers exist to scaling effective programmes or practices?

What a good analytical question looks like

A well-structured analytical question should be clear and focused. This means it should address: (1) the topic and scope of the analysis, (2) the target population or context, and (3) the purpose of the inquiry.

- **Topic and scope**: What issue is being addressed?

- **Target population and/or context**: Who or what is affected by this issue?

- **Purpose of inquiry**: What decision or action will the findings support?

Below we provide guidance on how to craft a strong analytical question.

Make the question clear and specific. Avoid questions that are too broad or overly general, as this c unmanageable range of evidence and vague conclusions. Omitting geographic, institutional, or der parameters can result in irrelevant findings.

Instead of...	Try...
What improves education outcomes?	What school-level factors contribute to improved outcomes for primary aged students in low-incom
What are the effects of poverty on education?	How does household poverty shape parental invo children's education in rural communities?

Make open-ended, rather than yes-no questions. This encourages exploration of multiple factors ar perspectives, resulting in a more nuanced understanding of the issue.

Instead of...	Try...
Does technology improve education outcomes?	Under what conditions does technology use in cla improve student engagement and achievement?
Are teacher recruitment policies effective?	What factors influence the effectiveness of teache policies in addressing shortages in crisis-affected

Don't frame it as a leading or biased question. Analytical questions should be neutral in order to cha assumptions and lead to new and meaningful insights rather than a predetermined outcome. A bia assumes an answer before the evidence is examined, making the research less objective or credible

Instead of...	Try...
Why are private schools better than public schools?	How do student performance and teacher suppor between public and private schools?
How do unqualified teachers negatively impact student learning?	What is the relationship between teacher qualifica student learning outcomes?

Framing an analytical question is a crucial step as it determines the purpose of the enquiry and informs all subsequent steps of the analysis. A well-structured question should be clear and focused, addressing specific issues, populations, and contexts. By

avoiding vague or overly general questions, researchers can ensure their findings are applicable and actionable. This approach enhances the overall decision-making process by providing targeted and relevant insights. The analytical question also helps identify the type of evidence sources that will be most relevant to finding the answers to the question. This topic will be explored in the next section.

Part 2. Identifying and accessing sources

What is the purpose of a systematic search plan

Our objective is to use the best available evidence for decision-making. There are many settings where evidence is in short supply. By broadening one's search to include research and non-research, published and unpublished sources, you are likely to find sources you were not aware of. However, this requires looking in the right places for the right sources. Because searching for grey literature sources is more complex than searching for published research, it is important to prioritise what you are looking for in terms of your analytical question, geographic focus, and timeframe. This section provides guidance on identifying and accessing relevant grey literature sources with the aim of being systematic and reducing bias in that sourcing.

What is the process for developing and implementing a systematic search plan

The process for developing and implementing a systematic search plan involves first identifying the types of sources needed to answer your analytical question (Step 1), then defining your additional search parameters (Step 2), and finally recording or tracking your searches, search outcomes, and challenges faced (Step 3).

Step 1. Identify types of sources to answer your analytical question

After creating an analytical question (Part 1), it is important to identify the most appropriate sources of information to answer that question. For example:

Example: If your interest is in how a programme is expected to be implemented and how it is being implemented, your search should seek:

- **Research sources**: Evaluations; Dissertations and theses; Technical Reports; Systematic reviews;

- **Government/Official sources**: Research reports; Guidelines; Memoranda/circulars

- **Organisational/Company sources**: Project/programme implementation reports; Evaluations; Manuals; Technical specifications and standards

As a first step, identify the types of sources that serve your analytical questions, drawing from Table 1 below. Table 1 is organised using the five classification clusters presented in the Introduction. It suggests where to search for each classification cluster and provides numerous examples, however, it is not comprehensive. This is just a starting point, and we encourage users to consult local education experts, librarians and other information specialists to identify more comprehensive sources of evidence depending on their needs and contexts.

Table 1. Examples of where to search by type of source

CLASSIFICATION CLUSTER	WHERE TO SEARCH	EXAMPLES
PRIMARY DATA	**Repositories and databases:** There are specialised repositories for original data	• Humanitarian Data Exchange https://data.humdata.org/dataset • UNESCO Institute for Statistics https://uis.u
	Contact authors Search for author's contact email through author's or organisation's website and write to request more information.	LinkedIn https://www.linkedin.com/ • Organisations' 'About Us' webpages • University department faculty webpages, e https://www.daystar. ac.ke/profiles/ and https://olpd.umn.edu/faculty
RESEARCH SOURCES	**Research maps, syntheses, and systematic reviews**	• American Institutes for Research (AIR) https://www.air.org/our- work/education • Best Evidence Encyclopaedia (BEE) https://bestevidence.org/about/ • The Campbell Collaboration https://www.campbellcollaboration. org/ • Centre for the Use of Evidence and Resear Education (CUREE) http://www.curee.co.uk/v • Education.org • Education Endowment Foundation https:// educationendowmentfoundation.org.uk/too
	Repositories and databases: There are specialised repositories and databases that focus on collecting and indexing grey literature sources and research. Search for local and regional journals that	• ADEA	Association for the Development of Africa adeanet.org • Africa Portal Digital Library https://africaportal.org/publications/ • African Education Research Database (AER https://essa-africa. org/AERD • African Journals Online (AJOL) https://www.ajol.info/index.php/ajol • Council for the Development of Social Scie Research in Africa (CODESRIA) https://codes • Open Alex https://openalex.org/

may not be indexed by international databases	• Database of Research on International Edu (ACER) https://opac.acer. edu.au/IDP_drie/ii • Early Childhood Development Action Netw https://ecdan.org/ • Education Resources Information Center (E https://eric.ed.gov/ • International Development Research Cent Africa Centre https:// idrc-crdi.ca/en/what- saharan-africa
Professional organisations, NGOs, and think tanks: Many professional associations, NGOs, and think tanks publish reports, working papers, and valuable research or analysis on various topics that may not be published through traditional channels. Visit their websites to search their publications to find relevant grey literature sources.	• Directory of Development Organizations http://www.devdir.org/ • Individual professional associations If a source is not available online, search for email through the organisation's website, pi association directories, LinkedIn, etc. and wi more information.
Expert networks and research communities: Engage with experts and researchers in your field of interest through professional networks, forums, or social media. They may be aware of unpublished studies or valuable sources that are not widely accessible. Some platforms require registration.	• Academic.edu https://www.academia.edu • Comparative and International Education such as CIES, UKFIET, BAICE, WCCES (see belo • Council of Ministers of the African and Mal Council for Higher Education (CAMES) https://www.lecames.org/ • LinkedIn https://www.linkedin.com/ • Education Evidence for Action (EE4A) https://educationevidence4action. com/ • Education International https://www.ei-ie.c • Educational Research Network for West an Africa (ERNWACA)

	Conference proceedings: Academic conferences often publish proceedings that include papers and presentations that have not yet undergone full peer review. Check conference websites or databases to access materials presented at conferences.	• AEN Evidence Conference https://africaevidencenetwork.org/en/ CIES https://cies.us/ • UKFIET https://www.ukfiet.org/
	Dissertations and theses: Graduate dissertations and theses, which can contain valuable analyses and other information.	• ProQuest Dissertations and Theses
	Library Catalogues	• National and University libraries, e.g. Daysto Kenya http:// repository.daystar.ac.ke/handle/123456789/
GOVERNMENT/ OFFICIAL SOURCES	**Government Publications:** Government agencies produce significant amounts of grey literature sources, including reports, policy papers, technical documents, and statistics. Check the websites of relevant government departments or search through government	• Global Partnership for Education https://globalpartnership.org/library • List of foreign government websites https://libguides.northwestern.edu/ ForeignGovernmentList • Ministry of Education and other governmer

databases to access this information.

ORGANISATIONAL/ COMPANY SOURCES	Organisational or institutional websites: Many organisations, including multilaterals, international non-governmental organisations, not-for-profits, research institutions, and companies, publish reports, working papers, and other literature on their websites. Some foundations provide open access to work they have funded. Explore these websites directly to access their publications.	• African Population and Health Research Centre https://aphrc.org/ publications/ • Bill and Melinda Gates Foundation https://gatesopenresearch.org/ • Echidna Giving https://echidnagiving.org/li • Forum for African Women Educationalists (https://fawe.org/ publications/ • Global Partnership for Education (GPE) https://globalpartnership.org/library • GPE KIX/IDRC https://www.gpekix.org/library • Jacobs Foundation https://jacobsfoundation.org/reports-studie • Open University https://www5.open.ac.uk/ikd/publications
INFORMAL AND OTHER ONLINE SOURCES	Discussion forums: Some valuable materials might be shared in online communities, forums, or discussion groups related to fields of interest.	• INEE Slack channel https://inee.org/comm practice
	Social Media: Follow researchers, practitioners, institutions, and organisations on social media platforms, as they may share links to	• Facebook • LinkedIn

grey literature
sources.

Step 2. Define key search parameters

The previous section helped you identify which sources are most relevant to answer your analytical question. Next you will need to define additional parameters to guide your search. It is important to think systematically about a search, recognising that it is not possible to be comprehensive in capturing all grey literature sources that may be available. Instead, your search parameters will determine what is included in your analysis, and what may be excluded, or out of scope. For example, ask yourself:

- **What websites, repositories, databases, and/or physical places will I go to find sources?** While Table 1 in the previous section provides an initial list to guide your planning, you may want to create a more comprehensive list, including local organisations, institutes, government offices, universities, or libraries.

- **What publication dates will I include?** You may determine that you want to look at studies from the past 5 or 10 years, or after a specific and relevant date (e.g. after a 2003 policy reform, since a new competency-based curriculum was launched in 2010, or after the Sustainable Development Goals were established in 2015).

- **In what languages will I accept sources?** You may decide to only include sources in one language, or to include multiple, local languages, including indigenous languages. Importantly, this decision should be determined based on the language capacities of your research team.

By being transparent and systematic about what you include or don't, it helps you to mitigate biases in the search process. Bias can be introduced in any search process. Mapping a search strategy and documenting the process will help mitigate biases introduced by depending on sources that are not relevant, credible, or representative.

Step 3. Record all searches, outcomes, and challenges

To ensure the search is systematic, it needs to be organised, with an accurate record of places searched and details on how the search was carried out. An example of this can be seen in Mitchell and Rose (2018). A search checklist template created in a spreadsheet, or a document table, can help keep the search focused. See Annex 2 for a Search Checklist Template.

It is recommended that, at a minimum, the following be tracked:

- **Date** the search was conducted

- **Name** of where you searched: database, repository, search engine, organisational website, or physical office.

- The **URL** of online platforms searched or **addresses** for physical offices.

If you need to outline your search methodology for an external audience, more extensive information will need to be recorded, including the filters used.

- **Geographic** coverage

- Search **terms** used

- **Date** ranges searched for

- Indicate search **outcome**, e.g. 'searched, nothing found'; 'searched, not relevant' 'searched, results found'; or 'searched, results may be of peripheral interest'

Keeping track of this information has several benefits. It will help maintain the focus of the search and ensure the necessary information is captured. Making it a transparent process, such as by adding search details in the write-up, helps to amplify the work of NGOs, CSOs, and researchers and highlights the value of such evidence. A comprehensive record of your searches and their results will also help better prepare you for subsequent follow-up studies on the same topic, saving yourself time in identifying relevant sources. Further, searching for grey literature comes with challenges (described in the box below). Documenting the challenges you encounter is also important, as it will help you later reflect on which places searched gave more relevant sources than others. Google may provide many, but how relevant were they compared to other types of searches? What sort of challenges did you encounter while searching and how can authors or publishers make their sources more easily accessible?

Challenges while searching for grey literature:

- Some organisations, databases, or repositories may have **paywalls or restricted access**. Others may have **limited search functionality**, making it more difficult and time-consuming to filter through and identify relevant materials. Annex 2 includes examples of several online repositories, demonstrating the distinct categories they use, or the different filters related to thematic area, type of document, date of publication, and more.

- The **use of different terminology** across organisations and sectors can result in inconsistent search results, making it critical to use a comprehensive list of search terms. Different terminology is often used to describe education levels (e.g. pre-primary vs. early childhood education vs. early childhood care and education vs. early childhood development) and education themes (e.g. inclusive education vs. special needs education vs. special education needs vs. disability inclusive education vs. disability inclusion in education). For example, in Education.org's synthesis on accelerated education, 26 terms used to refer to accelerated education globally were identified (e.g. accelerated learning, bridging programmes, re-entry programmes, etc.).

- **Incomplete metadata**, such as missing publication or production dates or unclear authorship, also complicates the process of identifying and citing sources. These are all challenges to accessing grey literature online, and they are exacerbated in contexts of low connectivity or where users may not have advanced digital skills.

- Finally, a lot of **grey literature may not be digitised**, making offline searches an important complementary effort. In some contexts, you may need to travel to libraries, universities or research institutes, as well as NGOs or other implementers to try and retrieve important evidence sources in person. Reaching out to your networks, and relying on personal connections takes time, but can be a productive way of accessing vital information.

Part 3. Appraising sources

What is the purpose of appraising sources

The process of appraising sources serves as a quality control mechanism. After identifying and collecting sources, it is important to use strong appraisal methods to ensure high-quality evidence. Traditional methods for appraising research are robust [1] but too limited for appraising non-research sources, like much of the grey literature sources covered by this Guidance. Additionally, other tools to appraise grey literature sources [2] do a good job in establishing source credibility but often miss key elements necessary for education decision-making.

This Guidance focuses on using evidence to bridge the gap between research and education policy, aiming to improve educational outcomes through evidence-based decisions. Therefore, this tool focuses on how well sources answer questions relevant to education decision-makers. This approach ensures that diverse types of evidence are considered in educational decisions. This Guidance builds on existing tools to offer a new, inclusive method for appraising individual sources for use. This tool is tailored to meet the needs of education decision-makers, focusing on how well sources answer relevant questions.

Introduction to the Appraisal Tool criteria

1. **Relevance:** These criteria look at the relevance of the source for the chosen analytical purpose or question, emphasising contextual relevance as well.

2. **Inclusivity:** These criteria look at the extent to which the perspectives of diverse and marginalised populations are included in the source.

3. **Credibility OR Methodological Rigour:** For all non-research sources and research sources that are non-empirical (e.g., evidence reviews and theoretical studies), these criteria look at the legitimacy of the source by exploring transparency dimensions and their accuracy. For research that is empirical (e.g., quantitative, qualitative, or mixed methods), the Mixed Methods Appraisal Tool (MMAT) is used to assess the source's methodological rigours.

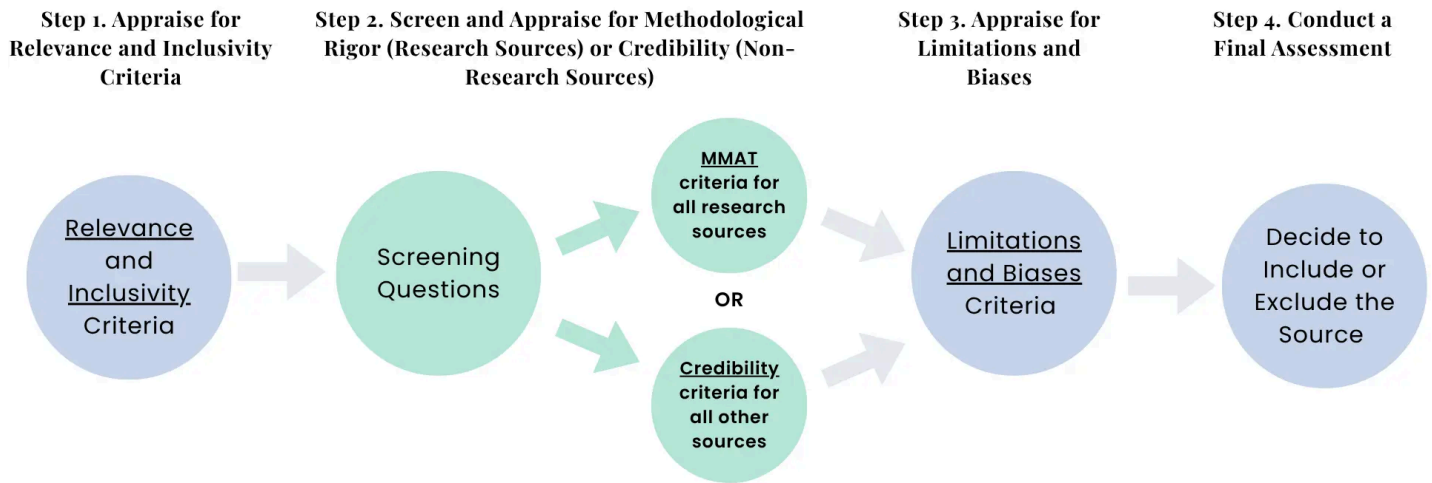4. **Limitations and Biases:** These criteria capture imitations and biases in the source

These four sets of criteria translate into the Appraisal Tool (Annex 3). Each criteria has sub-questions that are assessed as:

- **'Yes'** the source meets the criterion.

- **'No**,' the source does not meet the criterion.

- **'Unclear**,' there is no information available to answer the question.

- **'N/A,'** the question is not applicable for this source.

Users are asked to also include information describing why they selected their answer. This is to support their assessment of each criterion and is especially important if a 'No' or 'Unclear' assessment was determined. For the most part, a 'No' or 'Unclear' answer on a specific criterion is a red flag, meaning it may indicate limitations of a source's use, credibility, or diversity of voices, but does not necessarily mean that a source is unusable or not credible overall. However, it may raise questions and considerations, or require further investigation before use.

How to use the Appraisal Tool

The figure below describes the step-by-step process for appraising an individual source. Some criteria will be used to assess all sources (e.g. Relevance and Inclusivity in Step 1, Limitations and Biases in Step 3). However, during Step 2 you will answer two Screening Questions to determine whether the source is research or non-research. The Screening Questions will determine the criteria you use to appraise either Methodological Rigour (for research) or Credibility (for non-research). The final Step 4 is the same for all sources and asks to do a complete Final Assessment based on all the criteria. These criteria are explained in the following pages, and a template is provided in Annex 3.

Step 1. Appraise for Relevance and Inclusivity Criteria	Step 2. Screen and Appraise for Methodological Rigor (Research Sources) or Credibility (Non-Research Sources)	Step 3. Appraise for Limitations and Biases	Step 4. Conduct a Final Assessment

As indicated above, the criteria you use to appraise a source will depend on whether the source is research or non-research. The Appraisal Tool includes a set of Screening Questions (Step 2) to determine whether the source is research or non-research.

If the Screening Questions determine that the source is research, then the user is directed to use the Mixed Methods Appraisal Tool (MMAT) for appraising the methodological rigour of grey literature. We use the MMAT because it is the most versatile appraisal tool for using sources with a wide range of research methods. The MMAT includes appraisal questions for five different types of research methods: qualitative research, randomised controlled trials, non-randomised studies, quantitative descriptive studies, and mixed methods research. Using an appraisal tool that covers a wide range of research methods is thus essential when using grey literature sources.

Step 1. Appraise for relevance and inclusivity

Step 1 of the Appraisal Tool requires assessing a source against the Relevance and Inclusivity criteria. The sub-questions used to assess these criteria are found below in Table 2.

Table 2. Relevance and inclusivity appraisal criteria

APPRAISAL CRITERIA	GUIDING QUESTIONS (TO HELP EVALUATE EACH CRITERION)	EXAMPLES OF HOW TO AP
Relevance		
1. **Purpose:** Does the source's purpose or objective align with your analytical question?	Does the source clearly state or imply its purpose? Is this purpose aligned with your analytical question? Does the source fit the intended lens or perspectives required for your analysis?	A source with a clear and directly aligned with your score 'Yes.' If the purpose vague, or only partially re whether it scores 'Yes' or or not aligned with the to
2. **Relevance:** Does the source contain data, findings, or commentary that help answer your analytical question?	Is the source based on real experiences or events? Is the information current and appropriate for the intended use? Will the information in the source help answer your analytical question?	A source with direct, and score 'Yes.' A source with and some useful insights analytical question, may source is drawing on outc relevance to the analytic 'No.'
3. **Context:** Does the source address the needs, challenges, and circumstances of the context you are interested in?	Does the source clearly describe its context (for example, including setting, affected population, educational system, socio-cultural or political context, etc.)? Is the information comparable to or does it match the context of intended use?	A source with a clear, det aligned with your context If the context is stated bu partially relevant to your warrant either a 'Yes' or 'N not aligned with your focu
Inclusivity		
1. **Representation of marginalised groups:** Does the source focus on marginalised groups (students from low-income backgrounds, girls, refugees,	Does the source cover characteristics of marginalisation (e.g. rural or urban poverty, gender, disability, displacement, conflict, etc.)? Does it	A source that provides cle of the needs or challenge groups, or that analyses groups may sources that only look at

migrants, children or youth with disabilities, etc.)?	describe the needs, challenges, or support strategies related to marginalised communities?	population or groups of s differences based on ma
2. **Diversity of voices:** To what extent does the source acknowledge and involve a plurality of stakeholder perspectives and interests regarding the subject?	Are multiple perspectives captured or are the voices limited to one perspective? If the source was written by one person or institute (e.g. a government policy document), were other stakeholder groups and communities consulted in the process of developing it? If the source is research, is the sample of participants diverse?	Sources that score 'Yes' w perspectives of different s students, teachers, comn national authorities, or dif contexts (e.g. rural and u that focuses only on one city), or one type of stake leaders and not teachers

Step 2. Screen and appraise for credibility or methodological rigour

Step 2 of the Appraisal Tool requires using the Screening Questions to determine whether the source is research or non-research.

Table 3. Credibility or methodological rigour appraisal criteria

SCREENING QUESTIONS TO DETERMINE IF SOURCE IS RESEARCH OR NON-RESEARCH

1. Does the source address clear research questions or research objectives?	If answered 'Yes' to these questions, then the so it using the Methodological Rigour Criteria for R certain criteria based on the methodological d quantitative non-randomised, quantitative des questions, then apply the Credibility Criteria for
2. Is data collected to address those research questions or research objectives?	

Methodological rigour criteria for research sources

Research methodology	Appraisal criteria from the MMAT
Qualitative studies: sources with qualitative data collection and analysis, e.g. in-depth interviews or focus groups, case studies, ethnography, grounded theory.	Apply 5 MMAT Criteria: (i) Is the qualitative appr qualitative data collection methods adequate? the interpretation of results substantiated by th analysis, and interpretation?
Quantitative randomised studies: an experimental study in which participants are allocated to intervention or 'control' groups by randomisation.	Apply 5 MMAT Criteria: (i) is randomisation prop Are there complete outcome data? (iv) are out participants adhere to the assigned interventio
Quantitative non-randomised studies: any quantitative study estimating the effectiveness of an intervention without using randomisation to compare groups (e.g. non-randomised control trials, cohort studies, cross-sectional studies)	Apply 5 MMAT criteria: (i) are the participants re appropriate regarding both the outcome and i data? (iv) are the cofounders accounted for in intervention administered (or exposure occurre
Quantitative descriptive studies: studies used to describe quantitative variables without analysing causal relationships	Apply 5 criteria from the MMAT: (i) is the sampli sample representative of the target population response bias low? (iv) is the statistical analysis

| Mixed-methods studies: studies that involve combining qualitative and quantitative methods | Apply 15 MMAT criteria: the 5 appropriate quant (above), and the 5 mixed-methods criteria (i) i: (ii) are the different components of the study e the outputs of the integration of qualitative and divergences and inconsistencies between qua the different components of the study adhere t |

Credibility criteria for non-research sources

Appraisal criteria for non-research sources	Guiding questions (to help evaluate each criterion)	Example s of h
Credibility (for non-research sources)		
1. Author: Is an author (person or organisation) listed on the source?	If no author is listed, is the source produced, issued, or sponsored by an organisation, institution, or company?	If the source h should score 'N
2. Date: Is the year given for when this was written or published, or can it be estimated? Month is optional.	If no date given, can you estimate it either from references or an online search?	If the date of tl likely score 'Ye:
3. Transparency of source information: Does the source transparently document from where the information comes?	Does the author clearly explain where the information is coming from? Does the source cite its evidence, data, or origins of information? Are references, authors, or methodologies openly provided?	How this is dor with data tabl hearing or nev participant(s) one person's c
4. Transparency of description: Are the interventions, programmes, plans or policies referenced in the source clearly described?	Are the goals, targets, and activities of the programme, intervention, or policy clear? Are the processes, data sources, or analytical approaches	If the source h programme, o score 'No.'

used transparently
described?

| 5. **Soundness of argument:** Are the findings, conclusions, or recommendations supported by evidence? | What arguments does the author use to support the main points, arguments, or recommendations? Are the findings or conclusions the result of an analytical process, and is there logic in the points or opinions expressed? | A source that will likely score or thin, or the s without unpac |

Step 3. Appraise for limitations and biases

Step 3 of the Appraisal Tool requires assessing a source against the Limitations and Biases criteria. The sub-questions used to assess these criteria are found below in Table 4.

Table 4. Limitations and biases appraisal criteria

APPRAISAL CRITERIA	GUIDING QUESTIONS (TO HELP EVALUATE EACH CRIT
Limitations and biases	
Limitations: Does the source note any limitations of the work?	Is there a section labelled "limitations"? Are any go weaknesses discussed in the source? Limitations a methodological (e.g. small sample sizes, limited tir gaps in the data, etc.) or thematic (e.g. using one lens and not another).
Biases: Does the source have any evident bias, conflict of interest, or personal opinion?	Is the author's standpoint clear? If so, is it neutral c favour a particular perspective? Does the source p balanced analysis of the topic?

Step 4. Conduct a final assessment

The Appraisal Tool should be used to evaluate each individual source of evidence. **Step 4** of the Appraisal Tool asks for an overall Final Appraisal Evaluation. While it is up to each user (or team of users) to determine whether a source meets their unique needs, the aim of the Appraisal Tool is to support a consistent and structured assessment of sources. The criteria should help you in determining whether a source is relevant to your analytical question or country context, whether it draws on diverse perspectives, and whether it is credible or methodologically rigorous.

In the final assessment of the source overall, users will have to use their judgement to determine whether or not to include or exclude a source. To do this, look across all the criteria in the tool and indicate one of the final assessment options:

- **Include:** If there are few limitations identified (i.e. where the user selects 'Yes' for most criteria) then they may want to indicate 'Include.'

- **Include with reservations:** If most of the criteria are met with a 'Yes,' but perhaps some were 'No' or 'Unclear' pointing to gaps, weaknesses, or limitations.

- **Probably not include:** If most of the criteria are met with 'No' or 'Unclear' and there are only a few 'Positive' responses.

- **Exclude:** If almost all of the criteria are met with 'No.' This likely means the source is either irrelevant to your context, does not provide insights to answer your analytical question, or is full of biases and limitations that it may not constitute credible evidence.

Your final assessment does not necessarily determine what final decision you make. Importantly, you and your team should meet to discuss the sources reviewed. To minimise bias and enhance reliability, each source should be scored independently by at least two reviewers, followed by a comparison and discussion of results to reach consensus. A blank Appraisal Tool can be found in Annex 3, with space to explain your final assessment. If two reviewers assess a source as 'Not recommended for Inclusion' or 'Exclude,' the source should likely be excluded.

However, how your final decision on whether or not to include a source will also depend on how much evidence (or how many sources) you have access to in the first place. For example, you may end up choosing to include sources marked 'Include with reservations' or 'Probably not include' if they are the *only sources* providing the perspective of a specific stakeholder group (e.g. voices of teachers or students) or if they

are the only sources found covering a particular marginalised group (e.g. students with developmental disabilities). In other words, your decision on whether to include a source should be made while considering the entire evidence base, as this will be the information you use to answer your analytical question. It is important to note that using a wide range of types of sources can help complement and strengthen the analysis. As such, it is important to aim for a comprehensive analysis by appraising a range of sources.

Considerations for adapting the appraisal tool

The appraisal tool is designed to be adaptable, enabling diverse users (researchers and non-researchers) to assess the quality of a wide range of evidence, and different types of sources. Users should consider tailoring the application of the tool to their specific needs, evidence types, and thematic foci. Some practical ways to contextualise the tool are described below.

Develop clear indicators to apply each criterion: The Appraisal Tool should be applied consistently and objectively by each user within a given project or study. To help make this happen, consider adding specific indicators to guide users on what to look for when assessing each criterion on different source types. For example:

- "**Context**" will depend on the geographical and/or thematic focus of your research or analytical question. For example, if you are conducting a global mapping of interventions in education in emergencies, you will want to clearly define what constitutes an 'emergency.' You may decide that certain types of emergencies are more relevant for your work and should score higher on the appraisal tool. You may also consider emergencies in low- or middle-income countries as more relevant than those in high-income countries.

- "**Diversity of voices**" may look different for different sources: a source that is research may score high on this criterion if it includes a diverse sample population, including students, teachers, and school leaders from rural and urban schools. A policy document, on the other hand, may score high on this criterion if it was informed by consultations with different stakeholder groups. Your team may also decide to recognise diversity in authorships or publishers, for example, scoring those sources published by authors based in the local context higher than those sources written or published by individuals or organisations based in foreign or external contexts. For some research you may want to even divide the criterion into two: one to capture the research sample or content, and a second to capture the authors or researchers.

It is important that these indicators are clearly defined early on. You should test whether the tool is being applied consistently, by having different users assess the same source, and seeing if they come up with the same answer. Sample indicators for the MMAT tool can be found in Hong et al.'s (2018) <u>MMAT user guide</u>.

Adapt the scoring framework: To differentiate the quality of evidence more precisely, consider a scoring system that goes beyond binary options like "Yes" (1 point) and "No" or "Unclear" (0 points). For example:

- For "**Representation of marginalised groups**" you may want to divide this into two criteria: one that captures whether or not the source is authored by an actor from the focus context (e.g. to track authors from lower- and/or middle-income countries) and one that captures diversity within the content (e.g. whether the source includes the experiences of different marginalised groups, based on gender, disability, displacement, etc.).

- **Add a wider spectrum of answer choices.** You may want to elevate "Yes" to a value of 2 points and introduce a midpoint score of "Somewhat" (1 point). This "Somewhat" option would allow for partial fulfilment of criteria. For instance, for a criterion such as "Context," a global report with some sections addressing the specific needs of low-income countries, but limited examples from the Global South may score a "1" (Somewhat), while a report that focuses on the region of sub-Saharan Africa and provides detailed, country-specific data and analysis, may score a "2" (Yes). This nuanced approach enables users to rank evidence within a spectrum of quality, helping to identify those sources that meet the criteria more robustly.

Use the tool for prioritisation: When managing large volumes of evidence, such as in systematic reviews, the appraisal tool can also serve as a mechanism for prioritisation. High-scoring sources can be flagged for inclusion in deeper analysis or synthesis, while lower-scoring evidence can be deprioritised or used to highlight gaps.

Conclusion

This Guidance aims to make a methodological contribution to the education sector, serving as an approach that contributes to the use of the best available evidence in education decision-making. It is written for a wide range of actors who are producing and promoting the use of evidence in education policy and practice. It is hoped that it supports greater understanding of, and appreciation for, the wide range of education evidence sources beyond traditionally published research and elevate locally driven contextual insights that often are excluded through traditional sourcing approaches. The roadmap for identifying and accessing grey literature or non-academic sources, and the tool for appraising them, can lead to knowledge products that better serve education decision-makers and contribute to a 'world where the education of all children and young people is transformed by the best evidence' as envisaged by Education.org.

Users are encouraged to send their feedback on this Guidance to . This will help improve subsequent iterations of the Guidance.

1 Such as the Critical Appraisal Skills Programme (2018), Mixed Methods Appraisal Tool (Hong et al., 2018), and International Development Research Centre's Research Quality Plus Assessment Instrument (IDRC, 2022).

2 For example, Assessing unConventional Evidence (ACE) tool (Lewin et al., 2024), AACODS Checklist (Tyndall, 2010), and the Joanna Briggs Institute Critical Appraisal Checklist for Text and Opinion (McArthur et al., 2015).

Acknowledgements

Annex

International Working Group membership, alphabetical

(PDF)
Download ⤓

Search checklist template

(PDF)
Download ⤓

Examples of online repositories, filters, and key terms

(PDF)
Download ⤓

Appraisal Tool

(PDF)
Download ⤓

References

(PDF)
Download ⤓