VAKA YIKO

inasp

MODULE 3

ASSESSING EVIDENCE

MODULE 3

This trainer manual forms part of the VakaYiko Evidence-Informed Policy Making Toolkit. The Toolkit aims to support skills development and practical processes for evidence-informed policy making in public institutions in developing countries. It consists of a training course, a series of practical handbooks, and a range of informational and promotional materials.

This is the third in a four-part series of guidance notes for trainers. The complete Toolkit can be found on the INASP website here:

www.inasp.info/vytoolkit

Duration	Approx. 2 days and 3 hours [845–1,130 minutes]
Aim	To strengthen skills in critically appraising evidence for policy making.
Rationale	In this module, learners analyse and critically assess information from different sources.
Learning objectives	By the end of the module, learners will be able to: • judge the reliability and credibility of information found online, and apply this to their own work situations; • apply basic ways of appraising information (source, relevance, bias and quality); • compare and contrast the strengths and weaknesses of common evidence products; • assess a wide range of evidence products using checklists; • differentiate the elements of research design and use their understanding of these elements to help address quality.
Key learning points	• Four key considerations for assessing evidence for policy making are: source, objectivity, quality and relevance. • Each piece of evidence has different biases, strengths and weaknesses. There is no one 'perfect' product, so it is important to combine many different products to achieve a balanced view. • Understanding the basic elements of research design will help you assess the relevance and usefulness of different pieces of research for your topic.
Establish links	• Learners will work on critically assessing the evidence products they found in Module 2. • The various types of evidence and literature are still relevant here, as each type has different strengths and weaknesses.
Resources	• Projector and laptop for PowerPoint • Flipchart paper and different-coloured marker pens • Flipchart holders • Sticking tape • Small cards (exit cards) and post-it notes • Blank A4 paper

TOPIC 1 p.105

AN APPROACH TO CRITICALLY ASSESSING EVIDENCE [120–155 MINS]

ACTIVITIES:

M3-T1-A1	How to separate fact from fiction online	[40–50 mins]
M3-T1-A2	[Optional] How to fact check: tips and advice	[30–40 mins]
M3-T1-A3	Questions to critically assess evidence	[50–65 mins]

HANDOUTS:

M3-T1-H1	Questions to critically assess evidence

VIDEOS:

M3-T1-V1	How to separate fact and fiction online

TOPIC 2 p.112

ASSESSING EVIDENCE PRODUCTS [265–355 MINS]

ACTIVITIES:

M3-T2-A1	Assessing maternal mortality rates	[40–45 mins]
M3-T2-A2	Pros and cons of different evidence products	[45–50 mins]
M3-T2-A3	Pros and cons of work-related evidence products	[40–50 mins]
M3-T2-A4	Critically assessing different evidence products	[140–210 mins]

HANDOUTS:

M3-T2-H1	Assessing common evidence products
M3-T2-H2	Critically assessing different evidence products

TOPIC 3 p.118

UNDERSTANDING RESEARCH DESIGN [460–620 MINS]

ACTIVITIES:

M3-T3-A1	Experiences of empirical research	[20–30 mins]
M3-T3-A2	Research abstracts	[40–50 mins]
M3-T3-A3	[Optional] What is a clear and focused research question?	[30–40 mins]
M3-T3-A4	Categorizing research designs	[80–100 mins]
M3-T3-A5	Correlation does not imply causality!	[40–50 mins]
M3-T3-A6	Reviewing a research study design	[40–60 mins]
M3-T3-A7	Twenty tips for interpreting scientific claims	[50–60 mins]
M3-T3-A8	[Optional] Population and sample	[50–70 mins]
M3-T3-A9	[Optional] Sequencing and timing	[20–30 mins]
M3-T3-A10	Quantitative and qualitative research and methods	[30–40 mins]
M3-T3-A11	[Optional] External speaker presentation on research	[60–90 mins]

Optional Videos

HANDOUTS:

M3-T3-H1	Overview of the elements of research design
M3-T3-H2	Research abstracts
M3-T3-H3	Critical reading framework

Action plan and review activities (Trainer to build in)	• **Reflection on action plans** (to be carried out at flashpoints suggested throughout the course) [5–10 mins] • **Exit cards** (to be carried out at the end of each day) [5–10 mins] • **Review of Module 3** (to be carried out at the end of Module 3) [10–15 mins]
Further reading	**A critical view of systematic reviews for development policy:** ***www.odi.org/comment/6283-systematic-reviews-international-development-slrc*** **Africa Check** is an award-winning fact-checking organization which has lots of examples and guidance on verifying information on public policy issues: ***www.africacheck.org/*** **CLEAR – Regional Centres for Learning, Evaluation and Results** – strengthening capacities and systems for monitoring and evaluation (M&E) and performance management, to guide evidence-based development decisions: ***www.theclearinitiative.org/index.html*** **DFID – Assessing the Strength of Evidence: A How To Note**: ***www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf*** **'Is Your Evidence Robust Enough? Questions for Policymakers'** (Louise Shaxson, 2005): ***http://www.ingentaconnect.com/content/tpp/ep/2005/00000001/00000001/art00006*** **Systematic reviews and impact evaluations for international development topics** from 3ie: ***www.3ieimpact.org/***

TOPIC 1
AN APPROACH TO CRITICALLY ASSESSING EVIDENCE

MODULE 3 LEARNING OBJECTIVES RELEVANT TO TOPIC 1

By the end of this topic learners will be able to:

- Reflect on, and apply, basic ways of appraising information (source, objectivity, quality and relevance)
- Judge the reliability and credibility of information found online, and apply this to their own work situations

READ & REFLECT

This approach has been informed by Sutcliffe and Court (2005), Shaxson (2005), The Open University and Paul and Elder (2007).

The internet has completely transformed the way we access information and has made available a wealth of information that was previously difficult for the general public to access.

The flow of information, access to the internet and the increase in the number of open-access publications are all good things.

However, it also means we must be careful about the information we use; we can't rely on information in the same way as we could when everything went through formal review and publishing processes.

It is important to verify everything you read and to be careful about what information you choose to use as evidence to support decision-making.

The ability to critically assess what you read is a key skill for selecting the best evidence for informing decisions. The following approach can be applied to all of the four major types of evidence we consider (data, citizen evidence, practice-informed evidence and research).

WHY IS IT IMPORTANT TO ASSESS INFORMATION?

- Anyone can upload something to the internet;
- they can say anything they like – be it true or false;
- and leave it there as long as they like – even if it goes out of date;
- or change it without warning – perhaps even remove it completely.

THERE IS A DANGER THAT THE INFORMATION YOU FIND ON THE INTERNET MIGHT:

- be from a source that is unreliable,
- lacking in authority or credibility;
- have content that is invalid,
- inaccurate or out of date; or
- not be what it seems!

INASP, 2010.

To critically assess evidence, consider the following:

1. SOURCE AND CREDIBILITY

Identifying who provides the information is a key clue to its reliability. It represents the 'credentials' of a piece of information that support its status and perceived value. It is, therefore, very important to be able to identify the author, sponsoring body (i.e. the organization the author works in or that funded the research) or source of your information.

Factors to consider about authors:

- Are they acknowledged experts in the subject area? You could check this by doing a quick search through a search engine.
- Are they attached to a reputable institution?
- Have they been frequently cited by other authors in the field? In Google Scholar, for example, you can find out whether material has been frequently cited. Each search result shows how many times that study has been cited.
- Are they known to have a particular perspective on the topic? You could assess this by reading reviews of their work by other authors or the media, checking whether they have written further literature (such as opinion pieces) or participated in conferences.

Knowing good sources is helpful when time is scarce. For example, there are certain databases such as the Cochrane Library and the Campbell Collaboration that are reliable because the studies they provide go through a quality-checking process, which saves you time.

Factors to consider about sponsoring organizations:

- What type of organization is it: private company, NGO, research organization, policy institute, think tank, international organization?
- How well established is the organization? For example, how long has it been in existence? Does it work with reputable partners?
- Does the organization have any vested interests in the subject area being researched?
- How is the organization funded?

Factors to consider about the method of publication:

- Any individual can publish anything on the internet or post to a discussion list. This has to be judged on its own merit and with reference to the author's credentials.
- What do you know about the editor and/or the editorial board and how their editorial policy influences what will be published? Do you know if it has been submitted to peer review?
- Is the journal well regarded? Does it have a high rating in the Journal Citation Reports?

Remember that the source of a piece of information is not a direct clue to its quality. Sometimes renowned sources produce poor evidence, and little-known sources can also produce strong evidence. The 'stable theory' suggests that academic work is often valued highly just because it emanates from a prestigious research group or is published in a prestigious journal. We should judge information on its own merits.

This is an iterative process; with time, you will be able to build a trusted network of sources that you can refer to for different topics.

2. OBJECTIVITY VS. BIAS

In an ideal world, 'objective' or 'balanced' information would present all the evidence and all the arguments, and leave you to weigh this up and draw conclusions. In the real world, however, all information is presented from a position of interest. We also recognise that our own personal belief systems and opinions influence our ability to objectively evaluate information.

You will already have started to get some clues about bias and objectivity through your exploration of source and credibility. In some cases, authors may be expressing a particular viewpoint – this is perfectly valid as long as they are explicit about the perspective they represent. Hidden bias or errors of omission, whether or not deliberate, can be misleading, so it is considered good practice in formal research to clearly identify any potential biases. Therefore, your task is not necessarily to discount all biased information, but to ensure that you have identified the bias and allowed for it in your search process.

The familiarization exercises and contextual framework you created earlier should have provided you with some initial clues as to the potential areas of controversy in your topic, as well as the biggest stakeholders and what their interests might be.

Consider the following:

- **Information.** Is the evidence base clearly described, and are gaps/incomplete data acknowledged? Is it relevant to the question? Is there enough evidence? Is the information contrary to the conclusions included and explained?
- **Perspectives.** Do the authors state clearly the viewpoint they are taking? Do they situate themselves within current debates on the topic, identifying others whose position they agree with? Are different or competing viewpoints identified and addressed? Is there another way to look at the question?
- **Language** can be a useful danger sign. Look out for language that is either emotionally charged or vague. Assess the significance of the key concepts.
- **Sponsorship,** whether commercial, political or personal. For example, research may be sponsored by an industry or by a government. This does not necessarily make the research less objective, but it may make its interpretation selective. Make sure that all potential vested interests are clearly identified.

While no evidence product will be completely objective, combining the different points of view should mitigate these individual biases. If you fail to account for competing viewpoints, you run the risk of providing a biased answer to your research question.

3. QUALITY

Looking at the source of the information and assessing credibility and objectiveness should already have given you some strong indicators of quality. However, even credible, objective sources can sometimes produce poor-quality pieces of information!

Here are some questions to help you assess quality:

- **How was the information gathered?** We'll look more at detailed research methodology in Topic 4. But even with grey literature, data or programme reports from NGOs, you should be able to get a sense of how the information was gathered. Is there a methodology section? Was the information gathered through any kind of systematic process?
- **Does it include a range of types and sources of evidence?** Does it rely solely on one type or source of evidence? Is there a mix of qualitative and quantitative data? Are there any key types of evidence missing?
- **Do the claims made make sense based on the evidence presented?** A good piece of research will be very careful about what claims it makes. Researchers will usually use language such as "the survey results indicate that, within the specific population and context, X may be a contributing factor to Y". If your piece of evidence makes absolute or overblown claims about causality such as "X always causes Y" or "Z is the solution to Y", then you should proceed with caution.
- **Does it cite quality sources?** Regardless of whether you are reading a media article, blog, policy brief, PowerPoint presentation, academic article or any other source of evidence, always check whether the writer has cited their sources. This may be done in different ways, depending on the type of piece (a bibliography, footnotes, or citations within the text), but it is essential. If sources are not cited, this is an immediate indication that you cannot trust the quality of this piece of work. Look out for poor-quality citations such as Wikipedia, as well as sources that are out of date or unreliable.
- **What are other people saying about it?** When major think tanks and academic institutions release reports, other experts in the field often critically review these in the media and/or on blogs. Many academic journals also publish reviews of journal articles by other researchers. Reading critical reviews by other experts in the field can give you a sense of whether this piece of research is seen as credible and of high quality.
- **Timeliness.** Is it clear when the information was produced? Does the date of the information meet your requirements? Is it obsolete/ has it been superseded?
- **Language and presentation.** Is the piece well written? Are there any spelling or grammatical errors? Does it clearly indicate who the authors are, the date of production/ publication and, where relevant, the publisher and/or funder? Does it follow established format conventions for that type of evidence product?

Assessing the quality of research is not an exact science. But there are tools to help you do it, which we will explore more in Topic 4 on Understanding Research Design.

You will often find case studies or research that profile the right population you are looking for, but in a different context. Or you may find a study which focuses on exactly the right context, but the population is not quite the one you're looking at. Both can be relevant, but neither completely answers your question.

4. RELEVANCE

This is not a property of the research itself but the relationship between the research and your evidence needs. For example, you might be dealing with a piece of high-quality, objective information from a credible source but decide it is not relevant to the question you are asking or to the scope of your search.

The most important way to assess relevance is to be clear about *what question you are trying to answer* and *what type of evidence will help you answer it*.

The difficult part of assessing relevance is that you will rarely find a document which specifically answers your exact question. It is more likely that you will find a range of pieces of information which provide insights into different aspects of your topic.

For example, if you are looking to find out what is causing girls aged 15–17 to drop out of school in a certain town, you may not find much (or any) information on this exact topic. However, you might find:

- information on girls' high school drop-out in a range of regional contexts, including your own country, from global monitoring bodies such as UNESCO, but not focusing on the specific town you're looking at;
- survey results from an NGO working on school drop-out in the town you're looking at, but the data is not disaggregated, so you can't see how it specifically affects girls;
- other studies looking at youth issues such as work, early marriage and/or pregnancy, and lack of access to transport, which may mention school drop-out but don't focus on it specifically; or
- a study on your exact research question and target population but from a neighbouring country.

While on their own, none of these pieces of information completely answers your question, you can combine them to provide valuable insight into your issue.

Here are some questions you can ask yourself to help determine the relevance of a specific evidence product:

- How is the evidence connected to your evidence need and the type of question you are answering?
- Are there specific geographic limitations to the evidence you need (e.g. a study on cotton growth in arid soil in the western region of Zimbabwe)?
- If the evidence is from another context or with another population from the target you are looking for, is it a context/population which is still applicable to your situation, or is it too different to be useful?
- Are you seeking evidence about a specific period of time?
- Does the evidence address the complexity of the issue? Is it too complex or too simplistic?
- Are the findings/ recommendations widely applicable or context-specific?

KEY LEARNING POINT

Four key considerations for assessing evidence for policy making are: source, objectivity, quality and relevance.

REFLECTION POINT

In your workplace, how do you know if a source of information is of good quality? What do you do to check its credibility?

RECOMMENDED ACTIVITIES

PREPARATION

- Write up the learning objectives for the module on a flipchart and leave them displayed throughout the module so that they can be referred to at the start of each topic.
- In case of internet failure, print out the written transcript of the TED talk: ***www.ted.com/talks/markham_nolan_how_to_separate_fact_and_fiction_online/transcript?language=en*** for activity **M3-T1-A1**.
- Print out for each learner the cases in **M3-T1-H1. Questions to critically assess evidence** for activity **M3-T1-A3**.
- Write up questions for review activity **Exit cards** on a flipchart and label exit cards (three per learner).

M3-T1-A1

HOW TO SEPARATE FACT FROM FICTION ONLINE

[40–50 minutes]

1. Play the Markham Nolan video 'How to separate fact and fiction online': ***www.ted.com/talks/markham_nolan_how_to_separate_fact_and_fiction_online#t-85824***.
2. Invite the learners, while watching the video, to consider and make notes on the following:
   - three main ideas from the video that they think connect to the topic of this course/ workshop; and
   - any examples they have of misinformation resulting from reliance on poor information on the internet.
3. Invite learners to share and discuss their ideas and examples in plenary.

M3-T1-A2 *[OPTIONAL]*

HOW TO FACT CHECK: TIPS AND ADVICE

[30–40 minutes]

1. Organize the learners into groups of three and appoint a group leader.
2. Ask the group leader to decide how the group will read the Africa Check article 'How to Fact Check: Tips & Advice' (available online at ***www.africacheck.org/how-to-fact-check/tips-and-advice***). For example, each learner could read one or two specific parts or the whole group could speed read the whole article.
3. As a pre-reading task, ask the group to select and highlight up to three sentences which they find particularly interesting or useful.
4. Invite the groups in plenary to share their selected sentences together with their own examples or interpretations.

RECOMMENDED ACTIVITIES CONTINUED

M3-T1-A3

QUESTIONS TO CRITICALLY ASSESS EVIDENCE

[50–65 minutes]

1. Organize the learners into eight groups (with a minimum of two people) and hand out to each learner the cases in annex **M3-T1-H1. Questions to critically assess evidence**. Assign two groups Case A, two groups Case B, two groups Case C and two groups Case D. Ensure that the two groups working on the same case are not sitting close to each other.
2. If there are fewer than 16 learners in the group, then just divide them into groups so that all the cases in the handout are covered.
3. Check understanding of the task, and provide one or two examples from the Read & Reflect section only if necessary. Hand out two sheets of flipchart paper per group and marker pens for them to write down their questions. Ask learners to note down which case they are working on in a corner of their flipchart.
4. While the groups are working, note down which case each learner is working on. This list will help when forming new groups for activity M3-T2-A4 in Topic 2.
5. When the time is up, ask the groups to join the other group working on the same case. Invite them to share their questions, discuss and then agree on their final set of questions (between six and 12 questions as a guide) and ask them to write them clearly on one flipchart to present to the wider group. (If there are fewer than 16 learners in the group, then skip this step.)
6. Once the groups have prepared their flipcharts, ask each group to stick them on the wall or lay them on a table top so that the rest of the learners can move around the room and read them. Invite the learners to add any additional questions and/or make any comments on what they have read.
7. Review the flipcharts while learners are moving round the room and identify any gaps or missing questions to highlight (use the sample questions in the Read & Reflect section if needed). Keep hold of the flipcharts and type them up as soon as possible, as they will and used for activity M3-T2-A4.

EXIT CARDS

[5–10 minutes]

1. Carry out this activity at the end of each day.
2. Hand out the pre-prepared exit cards (three per learner) and ask each learner to write answers to the following three questions:
   A. What helped you learn today?
   B. What questions of clarification do you have/ areas you are unclear on from the sessions covered today?
   C. What comments or suggestions do you have for the trainers?
3. Gather the completed cards from the learners and explain that their comments will be reviewed after today's sessions and that there will be a short summary and response at the beginning of the following day's sessions.

TOPIC 2
ASSESSING EVIDENCE PRODUCTS

MODULE 3 LEARNING OBJECTIVES RELEVANT TO TOPIC 2

By the end of this topic learners will be able to:

- Compare and contrast the strengths and weaknesses of different evidence products
- Assess a wide range of evidence products using checklists

Each evidence product has different risks in terms of its quality. While research studies such as systematic reviews or formal peer-reviewed journal articles follow a process that enhances transparency and reduces the risk of bias, you still need to ensure that they are relevant to your question and purpose of searching for information. Some other products, such as research reports, policy briefs or think-tank working papers that are not submitted to quality and transparency processes, should go through a more detailed assessment to ensure you are selecting the most robust evidence available.

Remember that all the evidence products you are using will have different strengths and weaknesses, which is why it is important to combine several different products to answer your question.

TABLE 1
ASSESSING COMMON EVIDENCE PRODUCTS

Product	Type of literature	Pros	Cons
Systematic review	Peer reviewed; body of evidence	Covers a large body of evidence so can provide a valuable overview Checks the studies it reviews for academic rigour (bias, methodology etc.) so you don't have to Provides insight into different points of view/key debates in an issue Follows a transparent process that ensures scrutiny of the methods used and reduction of bias They increase confidence about what can be expected from an option (by increasing the number of units for study)	Usually written for an academic audience, so can be difficult to understand Can be very long Covers only peer-reviewed work (no grey literature) so may miss important types of evidence such as citizen evidence or practice-informed evidence Can be too general in scope to provide insight into very specific policy issues
Academic journal article	Peer reviewed (usually – do check!); can be single study or body of evidence	Produced through a thorough and rigorous academic process – very credible Builds on existing knowledge through literature review – the author has already considered the other academic literature out there Analyses primary data to help understand why and how things happen	Written for an academic audience, so can be difficult to digest No requirement to review other kinds of evidence aside from the academic (so may not have consulted much practice-informed evidence or citizen evidence) Can be too theoretical/remote, lacking concrete policy options
Annotated bibliography	Grey literature	Provides an overview of what has been written on the topic Short descriptions of evidence products, so you can choose which are relevant	It is a list only – doesn't synthesize/ summarize the research Can be difficult to tell from the short summaries what is useful to you
Policy brief	Grey literature	Specifically aimed at policymakers and focused on providing policy options Short, focused and engaging format; easy to read	Can be biased by specific interests, depending on source Not subject to formal academic quality checks (peer review, methodology etc.) No formal requirement to review existing evidence Cannot provide an in-depth picture
Programme report (e.g. from an NGO)	Grey literature	Rich source of practice-informed evidence providing insight into implementation Can be much more contextually/ geographically specific than other products	Can lack scientific rigour, as not a formal research product Can be biased (e.g. towards funders) No requirement to build on/review existing knowledge, so can tend to 're-invent the wheel'

Product	Type of literature	Pros	Cons
Report/paper from a think tank	Grey literature	A research-intensive form of grey literature which includes rigorous analysis Often more likely to consider a wide range of grey literature than an academic journal article would Aimed at informing policymakers; gives more thorough analysis than a policy brief	Can be biased depending on the ideological stance of the think tank
Statistics and data	Primary literature	Provides concrete quantitative information to provide a snapshot of a specific issue National statistics can provide information which is often not available in the academic literature International statistics (e.g. from WHO, World Bank etc.) are useful for comparison with other countries	The way the statistics are gathered/combined (i.e. quality of methodology) can be dubious Official government statistics may be unreliable, as some governments 'massage' the numbers for political reasons Cannot establish causality on their own – they don't explain why something happens Easy to misinterpret, not as simple as they seem
Impact evaluation	Grey literature; single study	Provides insights into a specific policy or programme to show the contribution of an intervention to a particular outcome Focused on real-life interventions in specific contexts and aimed at practitioners and policymakers rather than academics The most research-intensive form of grey literature – follows rigorous steps, subject to quality assurance processes and scrutiny by external specialists	Not usually subject to formal academic quality checks (peer review, methodology etc.) Can be very long

KEY LEARNING POINT

Each piece of evidence has different biases, strengths and weaknesses. There is no one 'perfect' product, so it is important to combine many different products to achieve a balanced view.

REFLECTION POINT

Choose a product that you are familiar with or you are using in your workplace. Explain why transparent processes and scrutiny of the methods used in an evidence product contribute to ensuring a high quality of this product.

RECOMMENDED ACTIVITIES

PREPARATION

- Print out hard copies of the article 'What is Zimbabwe's maternal mortality rate?' (available online at ***www.africacheck.org/reports/what-is-zimbabwes-real-maternal-mortality-rate***) for activity **M3-T2-A1**.
- Print out the table in **M3-T2-H1. Assessing common evidence products** so that there is enough for one handout per group for activity **M3-T2-A2**.
- Print out for each learner the Topic 2 Read & Reflect section for activity **M3-T2-A2**.
- Print out for each learner the cases in **M3-T2-H2. Critically assessing different evidence products** for activity **M3-T2-A4**.
- Depending on the size of the learner group, organise an additional breakout space for activity **M3-T2-A4**.
- Prepare document packages for activity **M3-T2-A4**. Depending on the level of the learners, choose two or three documents from a pool of documents of varying complexity stored in the electronic folder **M3-T2-A4 Document packages**. Select and print two or three documents from each of the four folders named objectivity, quality, relevance, and source and credibility. If it is a large group, print two sets rather than one set from each of the folders. Retrieve the final flipcharts with the learner questions from activity **M3-T1-A3. Questions to critically assess evidence**, type them up and print them out as one handout. Ensure that there are enough handouts to accompany each of the document packages printed from the obstacle course documents folder. Also add at a small pile of blank A4 paper to each of the document packages.
- Write up questions for review activity **Exit cards** on a flipchart and label exit cards (three per learner).

M3-T2-A1

ASSESSING MATERNAL MORTALITY RATES

[40–45 minutes]

1. Distribute to learners a hard copy of the article 'What is Zimbabwe's maternal mortality rate?'
2. Ask each learner to read the article and then discuss their initial thoughts and feelings about it in pairs or groups of three.
3. Invite the learners in the same groups to answer the following questions:
   - Who would have taken the DHS numbers? Why?
   - Who would have taken the census statistics, and who would have taken the more recent household survey figures? Why? (also UN modelling)
   - What can you practically do when all the data available is outdated?
4. Debrief the questions in plenary and ask learners for one or two pros and cons of this type of evidence.

M3-T2-A2

PROS AND CONS OF DIFFERENT EVIDENCE PRODUCTS

[45–50 minutes]

1. Organize learners into pairs or groups of three and hand out (one per group) the table in **M3-T2-H1. Assessing common evidence products** with the pros and cons columns to complete.
2. Invite the groups to brainstorm pros and cons for each evidence product and fill in the table (at least one pro and con per product).
3. Ask the groups to find at least two other groups to share and compare their tables with. Ask them to note down in their tables any new pros and cons they had not thought of.
4. Finally, ask the groups to return to their seats, hand out the Topic 2 Read & Reflect section and ask them to compare their completed tables with the one provided in the handout.
5. In plenary, invite three or four groups to briefly share any pros and cons they had not thought of but which were listed in the Read & Reflect table or vice versa.

M3-T2-A3

PROS AND CONS OF WORK-RELATED EVIDENCE PRODUCTS

[40–50 minutes]

1. Invite learners to refer back to the list of evidence/documents they searched for and compiled on their work-related topic at the end of Module 2 (activities M2-T5-A2 and M2-T5-A3).
2. Ask them to select three different evidence products (if possible) and assess them against the pros and cons table in Topic 2 Read & Reflect, ticking off which factors they find.

M3-T2-A4

CRITICALLY ASSESSING DIFFERENT EVIDENCE PRODUCTS

[140–210 minutes]

1. Place the objectivity document package/s, quality document package/s, relevance document package/s, and source and credibility document package/s separately around the room, and in a breakout space if one can be organised. Ensure that the document packages around the same assessment criterion are not placed near each other.
2. Invite the learners to move to a document station which focuses on an assessment criterion (i.e. objectivity, quality, relevance or source and credibility) they did not write questions for in the previous activity in Topic 1. Check the list of names noted during the questions activity **M3-T1-A3**, if necessary. Make sure that there is a reasonable balance of learners in each group across the different document stations.
3. Distribute to each learner the cases in **M3-T2-H2. Critically assessing different evidence products**. Ask the groups to choose and read the case that corresponds to the assessment criterion at the document station where they are sitting. Inform learners at the source and credibility document station/s that at least one of them will need to have a computer with internet access to successfully complete the task. If no internet is available at the time of the activity, then remove the source and credibility document packages and ask learners to also assess source and credibility along with the other three criteria (quality, relevance and objectivity).
4. Check that learners understand the task and invite the groups to prepare and bullet-point their assessment of the evidence products within the document package, on the blank A4 paper provided. Remind groups that they can use as an aide the typed list of questions they compiled in the question activity in Topic 1.
5. When the time is up or the groups feel they have finished, ask the groups to move to a different document station to work on a different case.
6. Depending on the time available and the level of engagement of the learners, stop the activity there or give them the option to move to another document station to work on a third case.
7. Decide on the most appropriate way for learners to provide feedback on the written assessments. For example, one group could present their main findings on one case in plenary, with the other groups and the trainer adding their additional points and/or comments. A second option is for groups who worked on the same case to swap their written bullet-point assessments with each other for immediate written and/or verbal feedback. A third option is for the groups to submit their written bullet-point assessments to the trainer for feedback via email or on paper.
8. Conclude the activity by highlighting the key points around the four criteria outlined in the Read & Reflect section.

RECOMMENDED ACTIVITIES CONTINUED

EXIT CARDS

[5–10 minutes]

1. Carry out this activity at the end of each day.
2. Hand out the pre-prepared exit cards (three per learner) and ask each learner to write answers to the following three questions:
   A. What helped you learn today?
   B. What questions of clarification do you have/ areas you are unclear on from the sessions covered today?
   C. What comments or suggestions do you have for the trainers?
3. Gather the completed cards from the learners and explain that their comments will be reviewed after today's sessions and that there will be a short summary and response at the beginning of the following day's sessions.

REFLECTION ON ACTION PLANS

[5–10 mins]

1. Display the slides again, if helpful as a reminder, in annex **M1ppt. Action plans**.
2. Invite learners to reflect on what has been covered in the course so far and write down notes under the key headings – i.e. challenges and ideas to support the use of evidence in policy making and to address the challenges identified.
3. Note that a longer session will be built in at the end of the course for learners to transfer their notes into the formal action plan. There will also be time to review their plans with the trainer and their peers.

TOPIC 3
UNDERSTANDING RESEARCH DESIGN

MODULE 3 LEARNING OBJECTIVES RELEVANT TO TOPIC 3

By the end of this topic learners will be able to:

- Differentiate elements of research design and use their understanding of these elements to help assess the quality and relevance of research

READ & REFLECT

The same general principles which we have discussed – source, objectivity, quality and relevance – also apply to research. Formalized and explicit procedures, standards and conventions, help to make empirical research objective, and the peer-review process of published journal articles functions as an important quality control mechanism.

As a user of evidence you do not only want to be able to make your own interpretation of results but also have an understanding of how the results came about – i.e. know whether the methodology and research design are appropriate to answer the research question.

The research design can be thought of as the method of planning research to gather the most appropriate information, in the correct way, and to analyse the results effectively (Laws, Harper, Jones and Marcus, 2013). It consists of different elements and acts as the 'blueprint' of a study. Each element will be explained in more detail in Topic 4. An overview can be found in Table 2.

WHAT'S THE DIFFERENCE BETWEEN EMPIRICAL AND THEORETICAL RESEARCH?

Empirical research aims at the development of new insights through the collection of data (empirical = observation or measurement rather than theoretical reasoning).

Theoretical research generally uses existing theories or hypotheses to develop new ideas. These new ideas are not tested by collecting evidence.

CORRELATION AND CAUSATION

One of the most common mistakes made when reading research papers is to conflate correlation with causation.

Correlation is the association between two variables such that when one changes, the other also changes. Correlation does not prove causation.

A causal relationship is a relationship between two or more factors in which one factor directly explains the other.

TABLE 2
OVERVIEW OF THE ELEMENTS OF RESEARCH DESIGN

Element	Explanation	Questions to consider
Research question	In natural and social science, the research question is the starting point of every study. Often this question is derived from theoretical considerations and implications or a gap in the literature. A good research question is sufficiently focused (e.g. determining location, type of research design, population and objectives of a study).	Is the research question explicitly stated? Where did the researcher derive his/her question from? Is the research question specific enough to guide the research? Is the research question answerable?
Type of research design	The type of research design refers to how research or studies may be categorized according to certain similarities and differences. Typical designs include experimental and longitudinal studies, among others. Each design has its advantages and disadvantages.	Is the research design type suitable to answer the research question? Does the research design type allow causal conclusions? Do other studies with different research designs reach the same conclusions?
Population and sample	The population and procedures to draw a sample are crucial to generalize findings of a particular piece of research. Ideally, the research draws a probabilistic sample or proves that the sample corresponds to the targeted population.	What is the study population? Is the study population of interest for my information need? Is the sampling procedure explicitly described?
Timing and sequencing	Research may investigate a problem or phenomenon over time (process), such as longitudinal studies, or at a particular point in time (state), such as cross-sectional studies.	Do findings from the research apply to different times?
Data collection methods	Typically data is collected qualitatively – for instance, through open questions in interviews – or quantitatively – for example, by measuring unemployment rates. Quantitative and qualitative data is analysed differently and has different advantages and disadvantages.	Is the data quantitative or qualitative? Is the quantitative indicator a valid measure of the phenomenon?

Ideally, the researcher would have defined each of the elements of the research design before the start of the study. This helps to reduce biased and unstructured data collection and analysis. As good practice, clinical studies are sometimes registered beforehand, and the researcher is held accountable according to this pre-defined and registered plan.[1] To publish in high-impact medical journals, this registration process is often a requirement.

1. See, for instance, ClinicalTrials.gov.

RESEARCH QUESTION

The research question is the starting point for a study. However, defining this question is by no means straightforward and usually involves an intensive review of existing literature and theories. This is because a researcher often aims to generate new insights.

A research question should be clear and focused enough to guide the research design. There is no specific standard about how a research question needs to be phrased, but it may contain some or all of following characteristics:

- **Answerablility.** For example, *"Are children good at mathematics?"* is not an answerable question.
- **Population.** Define who is being researched, referring to a population of individuals or objects. For example, *"How are the mathematical skills of school children aged 6 years?"*
- Set out **what** particular issues, events and/or characteristics are being researched. This may contain a definition of particular outcome and treatment variables. For example, *"How do mathematical skills in school children aged 6 years differ between private and public primary schools?"*
- Define the study's **timing and sequencing**, if it investigates a process or state. For example, *"How do mathematical skills develop in school children from age 6 to age 8?"*

The research question should appear in the first part of the study (or in the summary). You can check how relevant this research question is for your own information need.[2]

Once the research question is defined, the study's design can be planned more specifically.

TYPE OF RESEARCH DESIGN

The research design can be categorized into different types. While it is important to understand these categorizations, judging by type does not replace a critical reading of a particular study. As a critical reader of a research study, you will always have to ask whether the research design is appropriate to answer the research question.

From reading the abstracts provided in Topic 4, you will also see that many studies differ from the categories proposed in Table 2. Remember that no study design is perfect and that there are overlaps and different ways of combining designs.

Other ways of categorization may be:

- qualitative, quantitative and mixed-methods designs;
- exploratory and confirmatory approaches; and
- others?

DOES YOUR EVIDENCE SHOW HOW YOUR ISSUE AFFECTS BOTH WOMEN AND MEN?

Gender sensitive evidence should be used to inform all policies and programmes, so that gender issues are mainstreamed and policies are designed to enable equal opportunities for men and women. It's important that evidence looks at the impact of policies on both men and women, in order to ensure equal access to resources and opportunities. The collection, production, analysis and use of gender sensitive evidence can help policymakers begin to address these needs. This can include gender statistics, research evidence which looks at the impact of policies on both men and women, citizen evidence which includes equal representation of men and women, and practice-informed evidence of what has and hasn't worked in terms of gender equality.

Gender statistics are not just relevant for monitoring the status of women. They can also be used to shed light on specific issues relating to men, such as men's risk of accidents, or harmful use of tobacco and alcohol. Critical as they are to designing effective policies and programmes, the production of gender statistics presents significant challenges to national statistical systems and many data gaps exist, particularly in areas such as poverty, time use, violence against women, and the environment.

Source: United Nations Department of Economic and Social Affairs, 2014.

2. In particular, quantitative research usually involves another step, namely the deduction of several hypotheses from the research question and/or theory. However, often these hypotheses are not explicitly stated in an academic paper. If you want to find out more about research hypotheses, please look into our Further Resources.

TABLE 3
DIFFERENT TYPES OF RESEARCH DESIGNS AND THEIR ADVANTAGES AND DISADVANTAGES

EXPERIMENTAL

Type of literature	Pros	Cons
Experimental designs are used to estimate the causal effects of an intervention. They have two key main characteristics. First, they *manipulate a variable*, also called treatment, grouping or independent variable. For instance, the treatment group would receive a new drug, and the control group would receive a placebo pill. In this case, the manipulated variable would be the intake of a new drug vs. a placebo pill. Second, experimental designs *randomly* assign study subjects to at least two different groups – for instance, for every study subject a toss of a coin decides whether they are assigned to the treatment or the control group. Examples of experimental designs are randomized controlled trials (RCTs). RCTs are considered the 'gold standard' in health research, especially when it comes to the assessment of drug or treatment effectiveness. They are, however, also employed in other fields such as education, agriculture and development. Experimental designs use a quantitative measurement of variables, though they may also employ qualitative elements.	The study's groups can be considered identical with regard to known and unknown confounding variables due to the random assignment of study subjects. The groups only differ regarding the manipulated variable (e.g. new drug vs. placebo pill). Therefore, one can conclude with higher confidence that any differences in the outcome variables are caused by the manipulated variable rather than pre-existing differences between the two groups. For example, after 10 days of taking the drug, the treatment group showed significantly fewer symptoms (outcome variable) than the control group receiving the placebo. The reduced number of symptoms can be attributed to the new drug.	Results cannot be easily generalized – i.e. the treatment may not operate in the same way in another time or place or with another sample.

Measuring the impact of microfinance

In a microfinance study, a large Indian microfinance institution, Spandana, identified 104 low-income neighbourhoods in Hyderabad, India, which were potential locations to open a branch office. Prior to opening the branch offices, 52 neighbourhoods were randomly selected to have an office open in 2005 – this became the treatment group. The remaining 52 neighbourhoods remained 'control' (receiving an office in the following years). Households were then interviewed 15–18 months after the introduction of microfinance in the treatment areas.

Source: Banerjee et al., 2008.

QUASI-EXPERIMENTAL

Type of literature	Pros	Cons
Quasi-experimental designs also aim to estimate causality but are different from RCTs in one key aspect: *study subjects are not randomly assigned* to the different groups. Instead, treatment and control groups are built through natural groups (e.g. two different classes in a school or persons living in different districts of a country receive different treatment). Groups may also be built through self-selection (e.g. the first people registering for the skill training will receive it). Similarly to an RCT, both groups are compared with each other after regarding the outcome or dependent variables. But because the groups may not be identical before the study started, the researcher usually conducts a pre-assessment to statistically control for known confounding variables, such as individual motivation, gender, socio-economic status or the outcome variable itself. Quasi-experimental designs use a quantitative measurement of variables, though they may also employ qualitative elements.	It can be used in situations where the researcher wants to establish a counterfactual – i.e. what would happen without the treatment – but cannot randomly assign people to the different groups because this is unethical or not feasible. If confounding variables are known (e.g. from literature review), they can be controlled for statistically, and causal effects hence estimated.	Increased risk that the study groups are not identical with regard to known and unknown confounding variables (e.g. in one group there may be more motivated persons, more males or generally persons with fewer symptoms). This may have an effect on the study's outcome variables and thus prevent any causal conclusions. Unknown confounding variables cannot be controlled for and hence bias results. The same disadvantages as for experimental designs apply (see p. 121).

Effect of training on the clinical management of malaria by medical assistants in Ghana

Malaria accounts for over 40% of all outpatient consultations in Ghana. A common problem associated with its treatment with the drug chloroquine is over- and under-dosage, and a preference for the intramuscular route of administration. Inadequate treatment is an important factor in the selection of resistant strains of malaria parasites. To ensure the proper management of diseases at health centres, the Ministry of Health instituted an in-service training programme for medical assistants in 1987. The study evaluated the effect of this training on the clinical management of malaria using a quasi-experimental design. Three methods of data collection were used: prescription survey, assessment questionnaires and focus group discussions. The findings revealed that gains in knowledge following the training had deteriorated within a year. There was also a discrepancy between knowledge and practice of malaria treatment. This was shown by over- and under-dosing of chloroquine in children and adults, respectively. There was also overwhelming preference (85% of all cases) for injections and a high tendency towards polypharmacy (average of five drugs per visit). The motivating reasons for these were mainly socio-cultural and included patient demand and attitudes, prescriber self-interests and stereotypes and the daily practical challenges of the community. While paying greater attention to supervision of clinical work at health posts, consideration must be given to socio-cultural context of drug use in any such future training programmes if rational use of drugs is to be achieved.

Source: Ofori-Adjei and Arhinful, 1994.

OBSERVATIONAL

Type of literature	Pros	Cons
Similarly to experimental and quasi-experimental studies, observational designs aim to draw inferences about potential effects of a treatment or intervention (grouping variable). However, the *manipulation of the grouping variable is outside the researcher's control*. Instead, grouping may be done according to time (e.g. before and after an event) or participant characteristics (e.g. comparing women with men or different age cohorts). Examples of this are cross-sectional surveys and longitudinal and cohort or panel studies.[3] Observational designs use a quantitative measurement of variables, though they may also employ qualitative elements.	Observational studies can provide important information from the general population or community-level data to design more informative experimental studies. Statistical methods that match two groups according to certain variables are used to reduce the effect of confounding variables and hence provide insights into causality – for instance, the effect of certain drugs on a particular health outcome. Longitudinal, panel data and cohort studies in particular can provide valuable information about developments over time.	Claims about cause and effect have to be treated with caution. There might be variables not considered or unknown by the study that moderate or influence the outcomes covertly. Longitudinal studies 'lose' their study subjects over time. Panel data may rely on unreliable statistics. Apart from these problems, the same disadvantages as for experimental designs apply (see p. 121).

Growth and poverty reduction in Uganda, 1999-2000: panel data evidence

To explore factors underlying growth and poverty reduction in Africa while overcoming some of the limitations of cross-country analysis, this article uses micro-level survey and panel-data evidence from Uganda spanning 1992-2000. The high elasticity of both income growth and poverty reduction with respect to agricultural output (coffee) prices confirms the benefits from Uganda's decisive liberalization of output markets. It also suggests the importance of product diversification to protect poor households against price shocks and the potential of improvements in the cotton market to tackle persistent poverty in the north. The importance of improving access to basic education and health care emerges more clearly than in cross-country analysis, but benefits depend on complementary investments in electricity and other infrastructure, and reductions in civil strife.

Source: Deininger and Okidi, 2003.

3. To find out more about the particularities of each of the observational designs, see the Further Resources section.

META-STUDIES

Type of literature	Pros	Cons
Meta-studies collect information and aggregate data from existing studies to summarize research in a particular field, point out research gaps and generate more generalizable and robust findings. Meta-studies comprise literature reviews, *systematic reviews and meta-analyses* as an important part of systematic reviews (for a more thorough explanation, see Topic 3). Systematic reviews are much more rigorous than traditional literature reviews and are generally less biased.	Meta-studies are very good at reducing the complexity and breadth of research. Generalizations of findings regarding time and population may be possible.	Unpublished research and research written in languages other than English is often not included in the reviews. There are debates as to how far synthesis/aggregation of findings based on different contexts is practical and whether it distorts results ("comparing apples with pears"). "Garbage in, garbage out" means that studies with bad methodological quality included in the review may distort findings.

What is the impact of microfinance on poor people? A systematic review of evidence from sub-Saharan Africa

The study rigorously and systematically reviewed the evidence to identify the impacts of micro-credit and micro-savings on poor people in sub-Saharan Africa and tested a causal pathway to understand why these impacts occur. It found that micro-credit and micro-savings make some people poorer and not richer. Clients save more but also spend more. Health generally increases and, for some, so do access to food and nutrition. Impacts on education are varied, with limited evidence for positive effects and considerable evidence that micro-credit may be doing harm, reducing the education of clients' children. Micro-credit may empower some women, while both micro-credit and micro-savings improve clients' housing. There is little available evidence about the impact on job creation or social cohesion. Exploring the causal pathway for these impacts shows how clients' failure to increase their income, determined by external factors as well as how they spend their money, can lead them into further debt, unable to invest in savings and reliant on further cycles of credit. Successful increases in income, repayment of loans and the accumulation of financial wealth are all feasible, but the analysis shows how these are not always achieved.

What are the implications? Micro-savings may be a better model than micro-credit, both theoretically (because it does not require an increase in income to pay high interest rates, so implications of failure are not so high) and based on the currently available evidence. However, the evidence on micro-savings is small, and further rigorous evaluation is needed. In conclusion, micro-credit and micro-savings are doing harm, as well as good, to the lives of the poor people whom they purport to serve. Cautious implementation and further rigorous evaluation are required if these interventions are to alleviate rather than deepen poverty.

Source: Stewart et al., 2010.

QUALITATIVE[4]

Type of literature	Pros	Cons
This group of designs comprises a variety of studies that take a qualitative approach to research. In contrast to the aforementioned designs, *qualitative designs* are more often exploratory, descriptive and seek to understand real-world problems or relationships, rather than measuring and quantifying them. For example, research that explores a community's behaviours and attitudes with regard to the use of technology in farming could be helpful when designing an agricultural programme or preparing a quantitative survey. Some participatory designs, such as action research, actively involve the study subject in the research to achieve social transformation. Qualitative research is different from quantitative research in many ways. Please see the following section.	Qualitative studies can provide valuable insights to inform future quantitative research and make sense of existing findings. Research that is done in a participatory[5] manner usually achieves a higher uptake of its findings and can thus be more effective.	Claims about cause and effect as well as generalizations have to be treated with caution because the number of study subjects or cases is usually very small (e.g. single case studies).

Exploring empowerment and democracy in Zimbabwe

This case study argues that an 'informed and alert electorate' is essential for the establishment of democratic governance in Africa and for the continent's future economic growth. This need is evident in Zimbabwe. This paper tells the story of a small community-based organization in a remote part of Zimbabwe, which helped to raise political awareness and consciousness among a disadvantaged rural population.

Source: Conyers and Cumanzala, 2004.

Adapted from DFID, 2014.

4. See the following sections for more information on qualitative research.

5. Please note that the other aforementioned research design types could also be participatory – for example, several stakeholders (researcher, policymakers, practitioners and/or study subject) are involved in the design and oversight of the study.

POPULATION AND SAMPLE

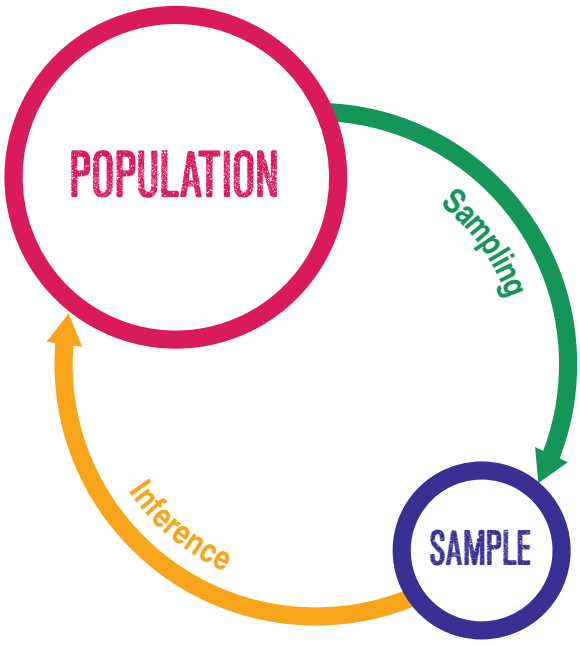
Usually in the abstract, executive summary or introduction of a piece of research you should be able to identify who the target population of the study was. Too often, we hear about a study being 'representative', and too rarely do we ask what this actually means.

Let's assume that an education researcher wants to find out about the mathematical skills of first-grade school children (study population), to identify where a mathematical skills programme should be conducted. They can make all first-grade school children take a maths test (census). This, however, could be a very resource-extensive study. Just imagine how many first-graders exist in your country and the logistical effort it costs to collect all the data. As an alternative, the researcher could ask a sample of first-graders to take the maths test. This would involve a lot less work for the researcher, and they can come to the same conclusions for the target population if the sample is drawn correctly – for example, through cluster sampling. Then, the researcher could infer from a sample of first-grade school children to the whole population of first-grade school children in the country. The mathematical skills programme could be adequately targeted.

There are different ways of drawing a sample – for instance, simple random sampling, cluster sampling, stratified sampling and so on – all of which can be termed *probability sampling*. Probability sampling means that each unit in the population has a chance of being selected in the sample. As a consequence, the sample has the same characteristics as the population, which allows generalizations. This stands in contrast to *non-probability sampling*, such as convenience, purposive and quota sampling, in which some units in the population have no chance of selection. The resulting sample does not necessarily have the same characteristics as the population, and generalizations are, therefore, often not possible.

Sampling is by no means an easy task. The sampling strategy, inferencing and/or extrapolation of results usually involves a skilled statistician, and adds to the quality of the study. You don't need to be an expert on sampling, but whichever study type you are referring to, make sure that the author clearly describes how the sampling was made. If a study claims representativeness, make sure you can identify the population referred to and that the sample has the same characteristics as the population.

FIGURE 1
SAMPLING AND INFERENCING IN RESEARCH



THE LAW OF LARGE NUMBERS

The Law of Large Numbers 'guarantees' stable long-term results for the averages of some random events. For example, you want know the average height of a group of 100 people. If you randomly choose one person, and measure their height, your estimate will likely be very far away from the population average.

However, if you randomly pick 25 people and measure their height, you will get very close to the true value in the population. It is important to remember that the Law of Large Numbers only applies (as the name indicates) when a large number of observations are considered.

SEQUENCING AND TIMING

This section is about the time and sequencing of conducting a study and how that relates to your own information need. Broadly, research may investigate a problem or phenomenon over time (process), such as longitudinal studies, or at a particular point in time (state), such as cross-sectional studies. In the following, different examples are given in which timing and sequencing play a crucial role, namely causality in longitudinal studies and panel data, pre- and follow-up tests and time of study.

TIMING FOR CAUSALITY: LONGITUDINAL STUDIES AND PANEL DATA

Longitudinal studies follow the study subject(s) over a defined period of time ranging from several days to a lifetime. They involve the regular measurement of the same sample over time, usually on the same variables, to identify changes. If you are interested in the development of characteristics or processes over time, make sure you look for longitudinal studies. These studies would answer questions such as disease pathogenesis or the development of mathematical skills in primary school children.

While the term *longitudinal studies* is used for biostatistics, econometrics often refer to multidimensional panel data instead. Panel data involves measurement over time, just like longitudinal studies. For their economic models, econometricians often use panel data based on indicators collected on a regular basis, sometimes over many decades, such as gross domestic product (GDP) or the human development index (HDI).

Causal links between different indicators should be well explained. For instance, it seems plausible to argue that economic growth over time (e.g. measured by GDP) has caused an increase in formal employment. But the other way around may also be true: that only with increased formal employment can the economy grow. In this example, we can make a causal link between increasing formal employment and GDP only if:

- GDP is increasing *after* formal employment has risen; and
- this effect is shown consistently over many time periods; and
- when alternative explanations have been ruled out.

Panel data or longitudinal studies can help to untangle this problem of causality on a macro level when RCTs are not suitable.

PRE- AND FOLLOW-UP TESTS

Experimental and quasi-experimental studies may involve a pre-test before an intervention takes place, to ensure that study groups have the same characteristics. They may also employ a follow-up, to assess the sustainability of effects in the medium and long term. Do not confuse a quasi-experimental study involving a pre- and follow-up assessment with a longitudinal study; longitudinal studies often work with natural groups (e.g. age cohorts), whereas quasi-experimental designs manipulate the intervening variable (see Table 2).

TIME OF THE STUDY

When reading a study, always check the time the data was collected. Findings from a study 50 years ago may not be relevant today, because, for instance, collective behaviours and norms change. In that case, you may want to look for more recent studies or commission a replication.

QUANTITATIVE AND QUALITATIVE RESEARCH AND METHODS

Data can be collected in different ways. One of the most fundamental distinctions is between quantitative and qualitative methods, which are each based on different research paradigms.

No matter which type of method has been used, it should be explicitly stated and documented in the methodology part of the research paper. See Table 4 for the advantages and disadvantages of qualitative and quantitative research. Where both qualitative and quantitative elements are used, this is called 'mixed-methods research'.

QUANTITATIVE RESEARCH AND METHODS

Quantitative research asks questions such as "How many?", "To what extent?" or "How much?" using counting and other computation. Quantitative research is concerned with the collection of data in the form of various measures and indices, and its description and analysis by means of statistical methods.

Quantitative methods produce numerical data, which comprises not just numbers such as height or weight but also different types of categories. Quantitative scientists speak of different data types such as categorical, ordinal and interval data and analyse these with the appropriate statistical methods, such as regressions, significance tests, correlations or simple counts and averages.

For example, if you count the number of people in the room and measure their height respectively, you would be able to:

- categorize learners into short and tall people (ordinal data) using, for instance, a median cut-off point;
- calculate the mean average height in cm (interval data); and
- relate from height to, for instance, gender by using correlation coefficients etc.

In public health studies, height could be used as an indicator or proxy for malnutrition in the development of children. So it is important that indicators are a valid reflection of the reality they seek to describe. Regarding measuring malnutrition in the development of children, body-mass index would be a better proxy than height, as it also accounts for body weight. Make sure you consider how well certain indicators reflect the reality they seek to describe.

Quantitative data can be collected in many different ways – for example, through rating scales or closed questions in questionnaires.

QUALITATIVE RESEARCH AND METHODS

Qualitative research describes the nature of answers (evidence) in terms of their verbal, written or other descriptive natures. It asks questions such as who, which, what, when, where and why? Qualitative research belongs to a family of approaches concerned with collecting in-depth data about human social experiences and contexts.

Returning to the example of malnutrition in children, a qualitative study would look at a few particular cases of malnourished children. It could look at contextual factors, such as unemployment in the family or climate change, and investigate the effects of the malnutrition, such as the resulting behaviour of the child or problems in the family. A qualitative research question could be: 'How does a family deal with a malnourished child?'

Non- or semi-standardized interviews, focus group discussions and observations produce a wealth of qualitative data in the form of interview transcripts and videos. Analysis is done according to different criteria based on methods such as hermeneutics, grounded theory and/or qualitative content analysis. It is difficult to generalize qualitative findings to a population because study subjects are often not selected randomly but according to their ability to contribute to the research question.

TABLE 4
ADVANTAGES AND DISADVANTAGES OF QUANTITATIVE AND QUALITATIVE METHODS, APPROACHES[6]

	Advantages	Disadvantages
Quantitative	The research results are relatively independent of the researcher (e.g. statistical significance) Can generalize a research finding when it has been replicated on many different populations and subpopulations Testing and validating already constructed theories about how and why phenomena occur Data collection using some quantitative methods is relatively quick Data analysis is relatively less time consuming (using statistical software) It is useful for studying large numbers of people	The researcher's categories and/or theories that are used might not reflect local constituencies' understandings The research may not be suited for explaining multiple aspects of complex situations Knowledge produced might be too abstract and general for direct application to specific local situations, contexts, and individuals Knowledge produced might not generalize to other people or other settings
Qualitative	Data based on the learners' own categories of meaning Useful for studying a limited number of cases in depth Useful for describing complex phenomena Can describe in rich detail phenomena as they are situated and embedded in local contexts The researcher almost always identifies contextual and setting factors as they relate to the phenomenon of interest Data is usually collected in naturalistic settings Qualitative approaches are especially responsive to local situations, conditions, and stakeholders' needs	Knowledge produced might not generalize to other people or other settings It is more difficult to test hypotheses and theories with large learner pools It generally takes more time to collect the data when compared to quantitative research The results are more easily influenced by the researcher's personal biases and idiosyncrasies

REFLECTION POINT

When do you use research information, and why?

KEY LEARNING POINT

Understanding the basic elements of research design will help you assess the relevance and usefulness of different pieces of research for your topic.

6. University of South Alabama, 2007a; 2007b.

RECOMMENDED ACTIVITIES

PREPARATION

- Print out for each learner handout **M3-T3-H1. Overview of the elements of research design** for activity **M3-T3-A1**.
- Print out for each learner handout **M3-T3-H2. Research abstracts** for activity **M3-T3-A2**.
- Prepare separate PPT slides for each of the different research designs in the following order: observational, quasi-experimental, experimental and meta-study for activity **M3-T3-A4**.
- Source and print out one or two examples (depending on the size of the group) of each of the five research designs (observational, quasi-experimental, experimental and meta-study) for activity **M3-T3-A6**.
- Depending on the size of the learner group, organize an additional breakout space for activity **M3-T3-A6**.
- Print out for each learner the article Sutherland, Spiegelhalter and Burgman (2013). 'Twenty tips for interpreting scientific claims'. Nature 503: 335, in the Readings and Samples list, and the table **M3-T3-H3. Critical reading framework** for reading activity **M3-T3-A7**.
- Prepare one or two PPT slides to support an explanation on the difference between probabilistic and non-probabilistic sampling, drawing on the 'Population and sample' section of the Read & Reflect section for **optional** activity **M3-T3-A8**.
- Prepare two PPT slides on longitudinal and panel designs, drawing on the Read & Reflect section for **optional** activity **M3-T3-A9**.
- Prepare a PPT slide or flipchart with definitions of quantitative and qualitative research methods for activity **M3-T3-A10**.
- For **optional** activity **M3-T3-A11**, invite a researcher to address the group on a piece of research they have been involved in, describing the different elements of the research design and why they chose them. It is important that the speaker is **prepared carefully in advance** so that they use the same terminology and draw on content relevant to this topic.
- Write up questions for review activity **Exit cards** on a flipchart and label exit cards (three per learner).

M3-T3-A1

EXPERIENCES OF EMPIRICAL RESEARCH

[20–30 minutes]

1. Ask learners to volunteer what they understand by the term 'empirical research'. Allow two or three learners to answer before providing them with a working definition.
2. Ask learners to organize themselves into groups of four and appoint a group spokesperson to provide feedback in plenary and a scribe to make notes during the group discussion.
3. Invite each group to *briefly* answer the questions below, focused on their experience of empirical research (stress to learners that the answers needs to be about research):
   - What research question did you want to answer?
   - What study design was used, and why?
   - What did the sample look like?
   - How was the data collected?
4. Invite each group spokesperson to present the answers from their group. Record the different research questions, study designs and data collection methods on three separate sheets of flipchart paper.
5. Ask learners to review the flipcharts and volunteer what they think are the key elements of a research design. Conclude the activity by handing out to each learner the table in annex **M3-T3-H1. Overview of the elements of research design** to read.

M3-T3-A2

RESEARCH ABSTRACTS

[40–50 minutes]

1. Hand out to each learner the research abstracts in **M3-T3-H2. Research abstracts** and ask each learner to read them and highlight with a pen the different elements of the research design they identify in each abstract. Note that not every abstract in the handout contains clues as to each of the design elements.
2. Invite the learners to share their answers in plenary.

RECOMMENDED ACTIVITIES CONTINUED

M3-T3-A3 *[OPTIONAL]*

WHAT IS A CLEAR AND FOCUSED RESEARCH QUESTION?

[30–40 minutes]

1. Ask learners to briefly answer, in pairs, the following two questions: "Why is it useful to have a research question?" and "What are the characteristics of a clear and focused research question in your view?" Invite two or three pairs to share their answers, before explaining its usefulness and the different criteria one could apply to it. Make reference to the search strategy from Module 2.
2. Organize the learners into groups of three or four and ask them to work on at least four of the different research questions provided in the previous activity and listed on the flipchart. Ask them to write down revised versions of the research questions listed and discuss how the question has improved.
3. If insufficient or no research questions were provided in the previous activity, ask the learners in their groups to note down some work-related research questions. Invite the groups to pass their research questions to another group to write down their improved versions.
4. In plenary, ask each group to share one or two examples of the 'before and after' research questions and ask the other groups to decide whether the question has improved or not and to explain why.

M3-T3-A4

CATEGORIZING RESEARCH DESIGNS

[80–100 minutes]

1. Refer back to the flipchart with the different research designs provided by the learners in the earlier activity. Ask in plenary whether anyone can suggest how they could be categorized in terms of different types of research designs. Depending on what learners come up with, explain the differences between the research designs and how they could be categorized (i.e. experimental, quasi-experimental, observational and meta-studies).
2. Give short presentations on the research designs, in the order listed below, followed by the discussion and recommended activity:
   - Explain *observational* designs, then ask learners to discuss in pairs the pros and cons of this type of design. Ask pairs for examples and for any questions of clarification. Go to activity **M3-T3-A5. Correlation does not imply causality!** opposite.
   - Explain *quasi-experimental* designs, then ask learners to discuss in pairs the pros and cons of this type of design. Ask pairs for examples and for any questions of clarification.
   - Explain what *experimental* designs are, then ask learners to discuss in pairs the pros and cons of this type of design. Ask pairs for examples and for any questions of clarification.
   - Explain what *meta-study* designs are, then ask learners to discuss in pairs the pros and cons of this type of design. Ask pairs for examples and for any questions of clarification.
3. If not mentioned already, ask learners whether they think all the research designs are of the same rigour and, if not, how they differ. If not already covered in discussion, ask learners what they think is the most common problem associated with the four study designs presented – i.e. that it applies a positivistic view to reality where everything is measureable.
   - Explain to learners, as an answer to the above shortcoming, that the last presentation will cover qualitative designs. Ask learners to discuss the pros and cons of this type of design, in plenary.

M3-T3-A5

CORRELATION DOES NOT IMPLY CAUSALITY!

[40–50 minutes]

1. Write up on a flipchart or whiteboard the words 'correlation', 'coincidence' and 'causality'. Ask learners to discuss in groups of three what the differences are between them, using policy-related examples where possible. Ask three or four pairs to share their thoughts in plenary, and encourage discussion among the wider group.
2. Ask each group to discuss and come up with three recent topical examples of causality and three topical examples of correlation.
3. Invite each group to present their examples in plenary, and encourage the learners to a lively debate over which are examples of causality and which are of correlation. Encourage the learners to question the rationale behind the choice of examples presented and to back up their positions by citing research and evidence.

M3-T3-A6

REVIEWING A RESEARCH STUDY DESIGN

[40–60 minutes]

1. Place one or two examples (depending on the size of the group) of each of the five research designs at different work stations around the training room, and the breakout space if available.
2. Ask learners to decide which type of research design they would like to focus on and to go to the appropriate work station. Some variance in group numbers will be acceptable but make sure there are at least three learners per work station.
3. Invite groups to familiarise themselves with the research study and discuss the questions in the handout **M3-T3-H1. Overview of the elements of research design**.
4. Ask groups to share their analysis in plenary and invite learners to volunteer any additional questions that could be used in their analysis of the research studies.

M3-T3-A7

TWENTY TIPS FOR INTERPRETING SCIENTIFIC CLAIMS

[50–60 minutes]

1. Ask learners to organize themselves into groups of three or four and appoint a group leader.
2. Hand out to each learner the article Sutherland, Spiegelhalter and Burgman (2013). 'Twenty tips for interpreting scientific claims'. Nature 503: 335, in the Readings and Samples list, and the table in handout **M3-T3-H3. Critical reading framework**.
3. Ask the group leader to decide how the group will read the article. For example, each learner could read four or more specific parts, or the whole group could speed read the whole article.
4. Explain the task and invite them to complete the table.
5. Invite the groups to share in plenary their selected quotes or ideas, together with their own reflections or interpretations. Recommend that the learners read the full article in their own time and consider the questions in the table.

M3-T3-A8 *[OPTIONAL]*

POPULATION AND SAMPLE

[50–70 minutes]

1. In plenary ask the learners why a researcher takes a sample. Explain the rationale behind taking a sample, using the diagram on slide four of the PPT in annex **M3ppt. Introduction and concepts**.
2. Check the learners' understanding of the terms 'sample', 'population' and 'sampling procedure', to find out how much they know about the topic. Briefly explain the terms if necessary.
3. Ask learners if anyone knows the terms 'probabilistic' and 'non-probabilistic sampling' and refer them to the handout **M3-T3-H1. Overview of the elements of research design**, which mentions one of the terms. Invite learners to explain what they think the terms mean and what the difference is between the two sampling methods.
4. Depending on how much learners know about the two sampling methods, give a short explanation of the difference between probabilistic and non-probabilistic sampling, using one or two PPT slides that draw on the 'Population and sample' section of Read & Reflect.
5. In plenary discuss and try out different ways (probabilistic and non-probabilistic) of taking a sample from the learner group (the population).
6. Ask the learners whether the resulting sample is representative of the group (use a visual characteristic) and ask them to reflect on PPT slide five on the Law of Large Numbers in annex **M3ppt. Introduction and concepts** in relation to sampling. Highlight any points not covered by the learners in discussion.
7. If considered useful or it will further develop the learning of the group, select one or both of the discussion topics below and ask learners in groups of three or four:
   - to discuss to what extent they think random sampling is different from randomization, as explained in experimental designs, and why each one is done; and
   - to identify the populations and sampling procedures in the different studies provided throughout this topic and discuss why samples are representative – or not – for the given populations – i.e. whether they have the same characteristics.
8. Invite the groups to share their key conclusions in plenary, and encourage discussion among the wider group.

RECOMMENDED ACTIVITIES CONTINUED

M3-T3-A9 *[OPTIONAL]*

SEQUENCING AND TIMING

[20–30 minutes]

1. Invite learners to consider whether they think research findings are generalizable across time. Explain longitudinal and panel designs using two PPT slides drawing on the Read & Reflect section.
2. In pairs, ask learners to discuss what kinds of topics/evidence requests would be best answered by longitudinal studies and panel designs.
3. Invite the pairs to share their ideas in plenary using examples from their work. Encourage discussion within the wider group.

M3-T3-A10

QUANTITATIVE AND QUALITATIVE RESEARCH AND METHODS

[30–40 minutes]

1. Ask learners to share in pairs what they know about quantitative and qualitative research methods. Invite the pairs to share their ideas in plenary.
2. Display a PPT slide or flipchart with definitions of each term. Ask learners to discuss the circumstances in which they find qualitative information useful, and when they find quantitative information useful. Encourage them to use real work-related examples.
3. Invite the pairs to join another pair to form groups of four, and to discuss and write down on flipchart paper the pros and cons of qualitative and quantitative research methods. Ask the groups to move around the room and add anything missing from the other groups' flipcharts. Explain, if not already mentioned, that both qualitative and quantitative methods can be used, which is called a 'mixed-methods' design.

M3-T3-A11 *[OPTIONAL]*

EXTERNAL SPEAKER PRESENTATION ON RESEARCH

[60–90 minutes]

1. An invited researcher makes a presentation to the group on a piece of research they have been involved in, describing the different elements of their research design and why they chose them.
2. In advance of the presentation, inform the learners of the title of the presentation and ask each learner to write down one question they would like answered in the presentation.
3. After the presentation, open the floor to the learners to ask the visiting researcher any of their questions that have been left unanswered.

OPTIONAL VIDEOS

'Worm wars': ***www.youtube.com/watch?t=188&v=9SCFIYINILQ***

Changing views on Zimbabwe's land reform: ***www.youtube.com/watch?t=3&v=t-7Vg0TNn2o***

Multidimensional Poverty Index: ***www.youtube.com/watch?v=yEULKXIokFw***

Richard Wilkinson – evidence on inequality: ***www.ted.com/talks/richard_wilkinson?language=en***

EXIT CARDS

[5–10 minutes]

1. Carry out this activity at the end of each day.
2. Hand out the pre-prepared exit cards (three per learner) and ask each learner to write answers to the following three questions:
   A. What helped you learn today?
   B. What questions of clarification do you have/ areas you are unclear on from the sessions covered today?
   C. What comments or suggestions do you have for the trainers?
3. Gather the completed cards from the learners and explain that their comments will be reviewed after today's sessions and that there will be a short summary and response at the beginning of the following day's sessions.

FURTHER READING

Richard Mallet – A critical view of systematic reviews for development policy: ***www.odi.org/comment/6283-systematic-reviews-international-development-slrc***

Africa Check – an award-winning fact-checking website: ***www.africacheck.org***

CLEAR – Regional Centres for Learning, Evaluation and Results – strengthening capacities and systems for monitoring and evaluation (M&E) and performance management, to guide evidence-based development decisions: ***www.theclearinitiative.org/index.html***

DFID – Assessing the Strength of Evidence: A How To Note: ***www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf***

Louise Shaxson – Is your Evidence Robust Enough? Questions for Policymakers (Policy Press, 2005)

Systematic reviews and impact evaluations for international development topics from 3ie: ***www.3ieimpact.org***

GLOSSARY

Annotated bibliography
A list of citations to books, articles, and documents. Each citation is followed by a brief (usually about 150 words) descriptive and evaluative paragraph, the 'annotation'. The purpose of the annotation is to inform the reader of the relevance, accuracy and quality of the sources cited.

Bias
Any influence which distorts or unduly influences the results of an investigation – perhaps as a result of the research method employed, sampling methods or the researcher's presuppositions. Some kinds of bias may be inevitable. To avoid any accusations of distortion, it is important to identify any factors you think may introduce bias (Laws, Harper, Jones and Marcus, 2013).

Correlation
The association between two variables such that when one changes, the other also changes. Correlation does not prove causality.

Causality
A causal relationship between two or more factors in which one factor directly explains the other.

Empirical research
aims at the development of new insights through the collection of data (empirical = observation or measurement rather than theoretical reasoning).

Experimental design
A research design in which the researcher tests the effects of an intervention by introducing the intervention to one group and compares this group with another which has not received the intervention (the 'control group').

Experimental evidence
Evidence as a result of an experimental study where a variable is manipulated and subjects are randomly assigned to a treatment or control group. For example, an intervention is given to one group of people but not to another one – i.e. the control group. The differences are then measured.

Evaluation
Aims to understand the performance and results from an organization, programme, project or any other intervention or initiative, using data captured during monitoring exercises conducted throughout the programme cycle.

Grey literature
Literature produced by government, academics, businesses, organizations and other institutions in formats not controlled by the commercial publishing industry.

Impact evaluation
An assessment of changes in the well-being of individuals, households, communities or firms that can be attributed to a particular project, programme or policy. The central impact evaluation question is what would have happened to those receiving the intervention if they had not in fact received the programme. Since we cannot observe this group both with and without the intervention, the key challenge is to develop a counterfactual – that is, a group which is as similar as possible (in observable and unobservable dimensions) to those receiving the intervention. This comparison allows for the establishment of definitive causality (World Bank, 2011).

Journal
A periodical in which articles relating to a particular discipline are published. Scholarly journals are often peer reviewed and present original research and reviews.

Literature review
A review of the current knowledge about a topic, including substantive findings, as well as theoretical and methodological contributions. A literature review is a standard part of any research paper (both formal academic papers and research reports from think tanks, NGOs etc). Although there are standard good practices for literature reviews, they do not follow as formal a process as systematic reviews and are not peer reviewed.

Non-experimental evidence
Evidence as a result of an observational study that describes research which observes (and explains) the effects of something already taking place in the real world. This means that – unlike experimental or quasi-experimental designs – the researcher does not directly design or implement the intervention themselves.

Observational design
A research design which observes the effects of something already taking place in the real world and where the researcher does not directly design or implement the intervention.

Policy brief
A short paper (usually three to four pages) that covers a specific issue. Typical briefs have four main functions: to explain and convey the importance of an issue or outline a problem; to present solutions and policy recommendations; to provide evidence to support the reasoning behind those recommendations; and to point the reader to additional resources on the issue.

Qualitative methods and data
The nature of answers (evidence) in terms of their verbal, written or other descriptive natures. It asks question such as who, which, what, when where and why? Qualitative research belongs to a family of approaches concerned with collecting in-depth data about human social experiences and contexts (Laws, Harper, Jones and Marcus, 2013).

Quantitative methods and data
asks questions such as "How many?", "To what extent?" or "How much?" using counting and other computation. Quantitative research is concerned with the collection of data in the form of various measures and indices, and its description and analysis by means of statistical methods (Laws, Harper, Jones and Marcus, 2013).

Quasi-experimental design
A research design which aims to measure the effects of an intervention, but without randomly assigning a group to treatment or control. These are often used when it is not practical or ethical to randomly assign people into groups.

Randomized controlled trial
A study in which people are allocated at random (by chance alone) to receive an intervention. One of these interventions is the standard of comparison or control. It can also be a type of impact evaluation which uses randomized access to social programmes as a means of limiting bias and generating an internally valid impact estimate.

Research design
The method of planning research to gather the most appropriate information, in the correct way, and to analyse the results effectively (Laws, Harper, Jones and Marcus, 2013).

Statistics: 'processed data'
It is the study of the process of collecting, analysing, interpreting, presenting and organizing data. Usually, governments have a unit or an agency that manages the statistics that have to do with their country. International/multilateral organizations such as the World Bank, WHO and African Union have useful statistical databases on their websites which enable comparison and analysis across countries.

Systematic review
A paper that gathers a large number of research papers and summarizes the findings through a specific, formal process which is peer reviewed. A systematic review uses transparent procedures to find, evaluate and synthesize the results of relevant research. Procedures are explicitly defined in advance, to ensure that the exercise is transparent and can be replicated. This practice is also designed to minimize bias. Studies included in a review are screened for quality, so that the findings of a large number of studies can be combined. Peer review is a key part of the process; qualified independent researchers control the author's methods and results (The Campbell Collaboration).

Theoretical research
generally uses existing theories or hypotheses to develop new ideas. These new ideas are not tested by collecting evidence.

Triangulation
looking at things from different points of view; the employment of a number of different research techniques, in the belief that a variety of approaches gives the best chance of achieving validity. This is because the way in which data is collected has an effect on the findings – for example, interviews, surveys, mapping exercises may all show different points of view on the same issue (Laws, Harper, Jones and Marcus, 2013).

Vested interest
A personal reason for involvement in an undertaking or situation, especially an expectation of financial or other gain – for example: "Banks have a vested interest in the growth of their customers" (Oxford Dictionaries, 2015).

REFERENCES

Banerjee, A., Duflo, E. Glennerster, R. and Kinnan, C. (2008). *Measuring the Impact of Microfinance*. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab. Retrieved 1 June 2016 from: ***www.betterevaluation.org/resources/examplemeasuring_impact_microfinance_using_rct***.

Conyers, D. and Cumanzala, F. (2004). Community Empowerment and Democracy in Zimbabwe: A Case Study from Binga District. *Social policy and administration*, 38(4): 383–398.

Deininger, K. and Okidi, J. (2003) 'Growth and Poverty Reduction in Uganda, 1999-2000: Panel Data Evidence' *Development Policy Review* 21(4): 481–509. Oxford, UK and Malden, MA: Overseas Development Institute and Blackwell Publishing. ***http://onlinelibrary.wiley.com/doi/10.1111/1467-7679.00220/epdf***.

Department for International Development (2014). *Use of evidence in policy making - scoping document*. London: DFID.

INASP (2010). *Science on the internet: a tutorial for parliamentary staff and policy advisors*. Oxford: INASP.

Laws, S., Harper, C., Jones, N. and Marcus, R. (2013). *Research for Development*. *A Practical Guide*. 2nd edition. London: SAGE.

Ofori-Adjei, D. and Arhinful, D.K. (1994). Effect of training on the clinical management of malaria by medical assistants in Ghana. *Social Science & Medicine*, 42(8): 1169–1176.

Oxford Dictionaries (2015). *Language matters*. Oxford: Oxford Dictionaries.

Parmesan, C. and Yohe, G. (2002). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421: 37–42.

Paul, R. and Elder, L. (2007). *The miniature guide to critical thinking. Tomales, CA: Foundation for Critical Thinking.*

Shaxson, L. (2005). Is your evidence robust enough?. Questions for policymakers and practitioners. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(1): 101–111.

Stewart, R., van Rooye,n C., Dickson, K., Majoro, M., de Wet, T. (2010). *What is the impact of microfinance on poor people? A systematic review of evidence from sub-Saharan Africa*. Technical report. London: EPPI-Centre, Social Science Research Unit, University of London.

Sutcliffe, S. and Court, J. (2005). *What is it? How does it work? What relevance for developing countries?* London: ODI.

The Campbell Collaboration (n.d.). What is a systematic review? Retrieved 1 June 2016 from: ***www.campbellcollaboration.org/what_is_a_systematic_review***

The Open University (n.d.). Information skills for researchers. Retrieved 1 June 2016 from: ***www.open.ac.uk/infoskills-researchers/evaluation-introduction.htm***

United Nations Department of Economic and Social Affairs (2014). `Using data to measure gender equality'. ***www.un.org/en/development/desa/news/gender/using-data-to-measure-gender-equality.html***

University of South Alabama (2007a). Strengths and Weaknesses of Qualitative Research, Table 14-01. Retrieved 1 June 2016 from: ***www.southalabama.edu/coe/bset/johnson/oh_master/Ch14/Tab14-01.pdf***

University of South Alabama (2007b). Strengths and Weaknesses of Qualitative Research, Table 14-02. Retrieved 1 June 2016: ***www.southalabama.edu/coe/bset/johnson/oh_master/Ch14/Tab14-02.pdf***

World Bank (2011). Impact evaluation: Overview. Retrieved 1 June 2016 from: ***http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,menuPK:384339~pagePK:162100~piPK:159310~theSitePK:384329,00.html***