



What is an evidence hierarchy?

Leaders, managers, policy and program staff, and practitioners need to feel confident that their decisions are based on the best available evidence. One important aspect of this is knowing what constitutes high-quality research evidence. An evidence hierarchy is a useful tool to help decision-makers quickly identify the best available research evidence of 'what works' in their field.

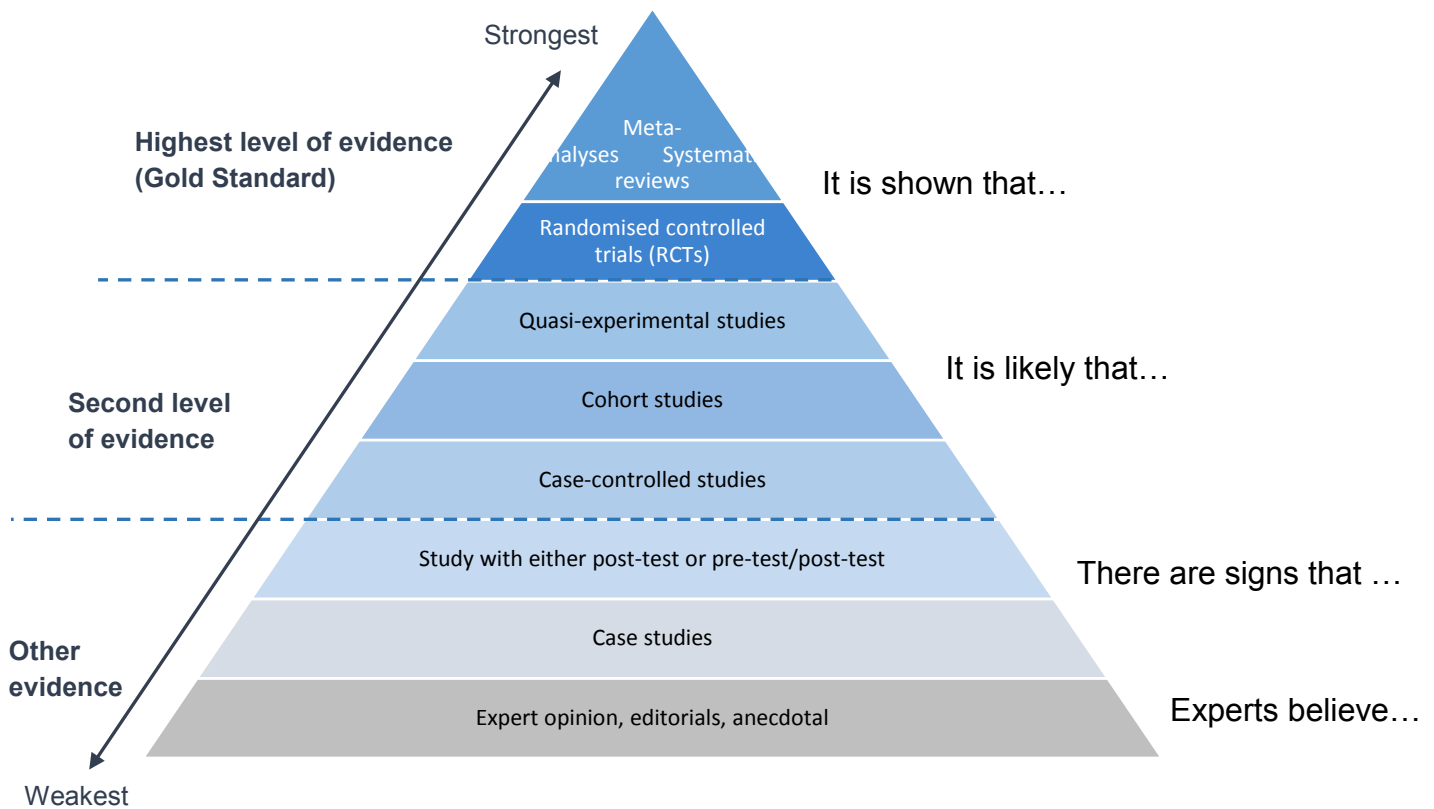
This Factsheet explains what an evidence hierarchy is and how it can be used to help you search for and appraise research evidence. It also provides links to useful resources.

The evidence hierarchy explained

Evidence hierarchies rank different research or evaluation study designs based on the rigour of their research methods. Evidence hierarchies vary, with different organisations using different criteria to group study designs. In general, the greater the number of high-quality studies included in the analysis and the more rigorous the research design, the higher the evidence rating. Research with the strongest indication of effectiveness, such as systematic reviews, meta-analyses and randomised controlled trials (RCTs), are usually at the top of evidence hierarchies.

Evidence hierarchies are often used in the health and medical field. They are less commonly used within social policy, as evidence with the highest indication of effectiveness such as RCTs and meta-analyses are rarely available. Where there is a lack of robust research evidence it is important to use a range of evidence types to draw conclusions and answer social policy questions. Figure 1 provides an example of an evidence hierarchy that has been adapted from the Centre for Education Statistics and Evaluation and National Health and Medical Research Council (NHMRC) classifications.

Figure 1. Evidence Hierarchy



Highest level of evidence (Gold standard evidence)


This category includes systematic reviews, meta-analyses and randomised control trials.

Systematic reviews

Systematic reviews are generally considered as the strongest form of evidence as they summarise and synthesise the findings of multiple studies identified in comprehensive, systematic literature searches. Systematic reviews that contain meta-analyses provide the most reliable summary of evidence on a topic.

Meta-analyses

Meta-analyses are based on the combined results of many studies rather than a single evaluation. The greater the number and diversity of participants and settings included in the analysis, the more convincing is the evidence.

Things to consider 

The quality of a systematic review depends on the quality of the studies included. A large, well conducted RCT may provide stronger evidence than a systematic review of small or poor quality studies. As systematic reviews can take a long time to produce it is also important to check when the literature search was conducted as there may be more recent evidence available.

Randomised control trials

Individual RCTs provide the next best possible evidence for assessing the impact of programs due to their robust research methodology. The key feature of an RCT is that participants are randomly assigned to either a 'treatment' group or a 'control' group – where the 'treatment' group receives the treatment or service under a program and the 'control' group does not. If the participants in the 'treatment' group have better outcomes than those in the 'control' group, it is likely that the program is effective.

Second level of evidence

This category includes quasi-experiments, cohort studies and case-controlled studies.

Studies in the second highest category of evidence have a comparison or control group but no random assignment. This means that program participants are deliberately matched with a similar comparison group rather than picked at random. These studies are more likely than RCTs to be inaccurate, as there may be some differences between the groups that cause a difference in outcomes. Careful matching and analysis to reduce the major differences between groups can sometimes increase confidence that outcomes can be attributed to the program.

Other evidence

This category includes post-test or pre and post-test studies of program participants (e.g. interviewing the participants before and after a program to see whether their opinions or behaviour have changed), case studies, qualitative research and expert opinion. Research studies in this category lack control or comparison groups, which means they cannot prove that a program is the cause of differences in participant's outcomes, or whether it was some other influence that was happening at the same time.

While evidence in this category cannot establish the effectiveness of programs, it can provide an indication of 'what might work' when there is a lack of more robust research evidence. This type of evidence is also important for answering other questions, such as why and how things are working.

Many program evaluations fall into this category, as often the only practical way to measure or evaluate a program's effectiveness is by comparing the outcomes of participants before (pre) and after (post). Although this is a good start to understanding the impacts of a program, a simple pre-post comparison should be treated cautiously as it is impossible to know if the changes in participant's outcomes were because of the program or other factors.

Case studies and qualitative research are usually smaller studies that involve just one, or a few people. Although they can provide valuable information about people's experiences they are not well suited to measuring program impacts because findings may not represent the entire population you are interested in. Assessments of effectiveness can also be influenced by the participant's memory (where data is collected after the program has finished) and there is a potential for bias in recalling information.

The opinion of a person regarded as an expert in the field is considered the least reliable form of evidence of effectiveness. However, anecdotal evidence may provide important

contextual information and can help to answer questions where no relevant studies are available.

Where can I find high-quality evidence of effectiveness?

There are several reputable sources of systematic reviews available:

- [The Campbell Collaboration](#) maintains a database of systematic reviews in the fields of education, social welfare, and criminal justice, with direct links to publications.
- [Cochrane](#) maintains a database of systematic reviews in the healthcare field that are peer-reviewed and can be downloaded from their website. The Developmental, Psychosocial and Learning Problems topic and Mental Health topic provides evidence that is relevant to the social care field.
- [EPPI-Centre](#) is a specialist centre for developing methods for systematic reviewing and synthesis of research evidence; and developing methods for the study of the use of research. EPPI-Centre maintains a database of systematic reviews that have been conducted by and supported by the EPPI-Centre.

Things to consider



Literature that has been ‘filtered’ or synthesised and critically appraised is attractive to policy makers and practitioners because it is often quick and easy to interpret and use in decision-making. However, different types of reviews vary dramatically in terms of quality and completeness.

Narrative reviews (often called Literature reviews) usually lack the systematic search protocols and explicit criteria for selecting and appraising evidence that are characteristic of Systematic reviews. This makes them very prone to bias. Most narrative reviews are opinion with selective illustrations from the literature, which do not qualify as adequate evidence to answer policy or practice questions. However, they can provide useful background information and an overview of the research on a topic.

Useful resources

The DCJ guide [Assessing the quality of research evidence](#) provides advice to help you develop a search strategy and critically appraise the evidence you find.

You may also want to refer to the [Centre for Education Statistics and Evaluation \(CESE\)](#) evidence hierarchy and the [National Health and Medical Research Council \(NHMRC\)](#) website for tips on rating the quality of evidence.

[What evidence should social policymakers use?](#) is an article produced by the Social Policy Division of the Australian Treasury that discusses the benefits of using an evidence hierarchy to rank evidence.

If you want to know more about RCTs, [Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials](#) is a helpful guide produced by the UK Cabinet Office.

Produced by

Strategy, Outcomes and Evidence
FACS Insights, Analysis and Research (FACSIAR)
NSW Department of Communities and Justice
320 Liverpool Rd, Ashfield NSW 2131
www.facs.nsw.gov.au
Email: facsiar@facs.nsw.gov.au

For more information contact FACSIAR at FACSIAR@facs.nsw.gov.au