What is an evidence-based program?

This Snapshot outlines some common definitions and classifications of evidence-based programs, drawn from international literature.

Introduction

There is growing emphasis in human services on making service delivery choices that are based on the best available evidence for what works, with program development and funding streams increasingly tied to research evidence.

Research evidence is one type of evidence that can be used to inform evidence-based decisions. Others include experiential evidence (professional expertise and insights) and contextual evidence (client and professional values).[1]

This Snapshot focuses on **evidence-based programs**, which in turn represent just one part of a body of research evidence. Research evidence is also available that relates to therapies, approaches, service systems and service models.

Definitions of an evidence-based program

The following definitions summarise some of the elements commonly found in evidence-based programs.

'Evidence-based programs refer to organized and typically multi-component interventions with clearly identified linkages between core components and expected outcomes, which have been experimentally evaluated and deemed efficacious in meeting specified goals.'[2]

'A discrete, organized package of practices, spelled out in guidance – sometimes called a manual or protocol – that explains what should be delivered to whom, when, where and how. A program is 'evidence based' when it has been evaluated robustly, typically by randomized controlled trial (RCT) or quasi-experimental design (QED) and found unequivocally to have a positive effect on one or more relevant child outcomes.'[3]
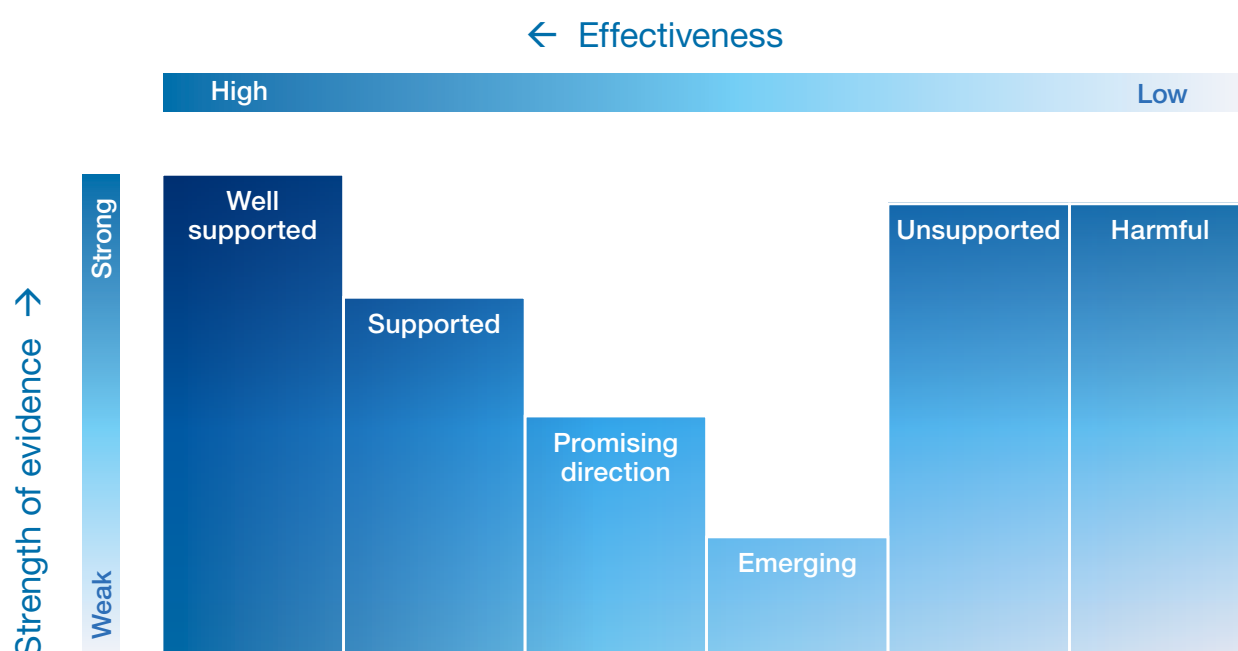
The determination of whether an individual program is evidence-based varies, with different organisations using different standards and rating scales to classify programs. Much of this variation is the result of different definitions of what constitutes rigorous research.

When is a program considered 'evidence-based'?

Evidence proving the effectiveness of a program can be thought of as falling on a continuum. The more rigorous the design of evaluations, and the greater the number of positive evaluations, the more compelling the evidence that the program will achieve its goals.

Most rating scales place programs along a continuum similar to the example at Figure 1, ranging from evidence of no impact through to well-supported or model programs.

Figure 1: Continuum of evidence strength and effectiveness

← Effectiveness

High ······· Low

Well supported

Supported

Promising direction

Emerging

Unsupported

Harmful

Strength of evidence → (Strong / Weak)

Adapted from Puddy & Wilkins (2011)

The **strength of evidence** underlying a program includes the rigour of the research design, and whether strong outcomes have been achieved in different settings with different populations (referred to as replication).

Another important consideration is the program's **effectiveness** – is it producing the desired outcomes, or leading to any undesirable outcomes?

Which research designs are considered the most rigorous?

In general, to be labelled 'evidence-based' a program has to have been shown to deliver significant and sustained positive outcomes. In addition, these positive effects must have been demonstrated through rigorous evaluation, with a large number of diverse participants, or through multiple replications. In most classification systems, this means that evidence of effectiveness must be demonstrated through multiple RCTs.

RCTs are considered the 'gold standard' of research design. The distinguishing feature of an RCT is that participants are randomly assigned to either a control group or a treatment group, enabling greater control over factors that may influence outcomes. If the participants of a program have better outcomes than those in the control group (who were not participants in the program), it is likely that the program is effective.

In a QED there is no random assignment – program participants are deliberately matched with a similar comparison group. Careful matching and analysis to reduce the major differences between groups increases confidence that outcomes can be attributed to the program.

Research designs that lack control or comparison groups cannot rule out other influences to prove that a program is the cause of differences in participant's outcomes. Although they may provide support for program effectiveness, these evaluations do not provide sufficient evidence for a program to be classified as evidence-based.

Programs that have insufficient evidence to determine their effectiveness are often referred to as emerging, promising or research-based.

Some classification frameworks go beyond assessments of program impacts to include criteria for quality implementation. These include a clear description of how the program operates, the theory of change (how components of the program work to produce outcomes) and a detailed implementation guide.

Other considerations

Evidence-based programs are not a 'one-size-fits-all' solution. When making decisions about the funding and implementation of services, it is important to consider other factors in addition to research evidence. These include:

Feasibility	the success of an evidence based program will be impacted by available resources, the economic, social, geographic and historical factors of the current setting, and organisational readiness and infrastructure.
Acceptability	to the people and decision-makers in the current setting.
Utility	how useful and appropriate is the EBP is to the target population and delivering organisation?[4,5]

Agencies considering implementing an evidence-based program should conduct a detailed examination of the implementation factors in relation to the service context, including factoring in the cost of implementation. Consideration should be given to the extent to which an intervention could or should be adapted to fit the context.[6]

Useful resources

Below are links to useful resources, evidence-based program directories, and further information about standards and rating scales used to classify programs.

Australian Research Alliance for Children and Youth (ARACY), Nest: what works for kids

Early Intervention Foundation: interactive guidebook on early intervention

California Evidence-Based Clearinghouse for Child Welfare

Blueprints for Healthy Youth Development

US Department of Education: What Works Clearinghouse

The PEW Charitable Trusts: Results First Clearinghouse Database

Endnotes

[1] Puddy, RW & Wilkins, N 2011, Understanding Evidence Part 1: Best Available Research Evidence. A Guide to the Continuum of Evidence of Effectiveness, Atlanta, GA: Centers for Disease Control and Prevention.

[2] Metz, AJR, Espiritu, R & Moore, KA 2007, What is Evidence-Based Practice? Child Trends Research-to-Results Brief, Washington, DC: Child Trends.

[3] Axford, N & Morpeth, L 2013, 'Evidence-based programs in children's services: A critical appraisal', Children and Youth Services Review, vol. 35, pp 268-277.

[4] Macvean, M, Sartore, G, Mildon, R, Shlonsky, A, Majika, C, Albers, B, Falkiner, J, Pourliakas, A, & Devine, D 2015, Effective Intensive Family Services Review, Report prepared by the Parenting Research Centre and The University of Melbourne on behalf of NSW Department of Family and Community Services.

[5] Puddy & Wilkins 2011.

[6] Macvean et al, 2015.