TEACH

What is Good — EVIDENCE?

March 2023

DOI: 10.26183/a6ph-np88

ozchild

WESTERN SYDNEY UNIVERSITY

Research Team

Professor Rebekah Grace
Tuli Keidar
Dr Cris Townley
Associate Professor Ann Dadich

ozchild

TEACH

Table of Contents

INTRODUCTION

In 2022, OzChild commissioned the Transforming early Education and Child Health (TeEACH) Research Centre to support them in their efforts to ground their programs and services on a solid foundation of evidence. OzChild is strengthening its organisational culture in which research both guides and responds to service decisions, practices, and policies. This report summarises the key issues to support an organisation-wide understanding of research evidence, the forms it can take, and how it can be used to support quality service provision.

Background and aim

Like many contemporary service organisations, OzChild has had their own 'journey to evidence' (McCarthy & Griffiths, 2021) as it has worked to ensure that its programs and supports are effective in achieving positive outcomes for children and their families. One approach to supporting evidence-based practice is the adoption and delivery of rigorously trialled and manualised evidence-based programs (EBPs), and OzChild has several programs of this nature (e.g., Treatment Foster Care Oregon, Functional Family Therapy, SafeCare, etc.). However, adopting an existing EBP is not always possible or appropriate, and so a large number of OzChild's current services and programs are 'bespoke' in that they consist of various elements of EBPs and evidence-based practices (e.g., Family Worx), have been designed to address place-based or cultural needs (e.g., Koorie Early Years Network), provide essential service provision (e.g., Family Law Services), or are large umbrella programs supporting cohorts of children with particular care needs (e.g., Foster Care, Kinship Care).

In recent years, there has been a shift in the way that researchers and practitioners have thought about evidence, from a rigid commitment to the strict implementation of rigorously assessed EBPs only, to an understanding of the importance of program adaption in response to local context and the importance of flexibility to address challenges as they arise. While the implementation of a well-developed program with an existing rigorous evidence base remains the gold standard, it is not possible if:

1. There are no 'off-the-shelf' EBPs that are aligned with the needs of a particular cohort or community;
2. A potentially applicable EBP was developed and assessed in a context or with a population group that differs in critical ways from the local context or population, throwing into question the validity of the research findings for the current context.

In such cases, aligning programs and practices to evidence requires directly engaging with the published research. This typically begins with searches for systematic reviews that examine multiple relevant interventions or that highlight the risk and protective factors that can be targeted by the program. However, delving into the literature can be daunting, particularly when undertaken without a clear framework or heuristic to interpret and appraise the evidence that is encountered.

Comparing the relative value, quality, and strength of different types of evidence is not straightforward. Although randomised controlled trials (RCTs), focus groups, and observational studies all produce valid forms of evidence, they are not all equally suited to answer the wide range of questions that are of interest to human services organisations. Accordingly, using evidence wisely – be it for program selection, implementation, or adaptation – requires a clear understanding of the question and the forms of evidence that provide the most compelling, reliable, and valid answers to that question.

The aim of this report is to provide OzChild with a practical guide on how to engage with research evidence in the assessment of their services. This in turn can support a broad understanding of the different forms of evidence and levels of rigour, and how research evidence can be meaningfully embedded into practice.

From theory to application

To evaluate evidence proficiently, one must first possess a nuanced understanding of what evidence is. In this paper, in the section Understanding Evidence, we unpack the concept of evidence, looking at how it is defined and applied in the human services sector. We then shift our focus to research evidence. We examine different philosophies of knowledge and describe the strengths and weaknesses of different research methodologies.

With a clearer understanding of evidence, we move onto Evaluating Evidence. Here, we discuss how evidence is evaluated by program clearinghouses, governments, and other actors in the human services sector, and interrogate how they reflect certain implicit definitions of evidence. We then present a more expansive view towards evaluating evidence, depicted through an evidence evaluation matrix.

The Using Evidence section delves into different ways that evidence can be used to guide program selection, implementation, and development. Before gathering or applying evidence, you first need to work out what evidence you need. For this, we introduce The Heptagon Tool – a handy method to structure program and implementation site evaluation. We then discuss the 'common elements' approach to program development to balance flexibility with adherence to practices drawn from solid evidence. Finally, we discuss the many benefits of embedding research into practice.

UNDERSTANDING EVIDENCE

Defining evidence

Evidence is 'the available body of facts or information indicating whether a belief or proposition is true or valid' (Oxford Languages, n.d.).This is a very broad definition, but for good reason: depending on the context in which evidence is used, its meaning changes. Legal definitions of evidence share little in common with how evidence is understood in a physics laboratory, or how it is thought of in the human services sector.

However, these definitions share the notion that facts or information (be they witness statements, survey data, or instrument measurements) only become evidence once they are used to answer a question, make a claim, or affirm a belief. The key point is that a piece of evidence can only be understood in relation to the question it answers, or the claim it supports. So long as there is integrity and transparency in its presentation, there is no such thing as inherently 'good' or 'bad' evidence–it all depends on the question.

Research as Evidence in Human Services

It goes without saying that the questions and claims posed in a courtroom differ from those that are of interest to service providers. Accordingly, the facts or information that count as credible and useful evidence also differ. In this report, our discussion of evidence will be limited to information acquired through research. Although there is an ongoing debate in the human services sector regarding what qualifies as credible evidence, it mainly centres around what types of research evidence should be privileged above others (Donaldson et al., 2009; Gray et al., 2009; Nutley et al., 2013). This is not to say that research is the only source of valid and useful knowledge: practitioner experience, expert opinion, and client voice are all vitally important, but they will not be considered evidence here except for when they are captured through research.

Research is granted the label of 'evidence' because researchers use systematic, rigorous, and transparent processes to gather, analyse, and publish data (Nutley et al., 2013). These processes are designed to maximise accuracy and reliability while minimising bias. Furthermore, when research is published in scholarly journals, the research methods and conclusions are scrutinised by other researchers during the peer-review process. This helps ensure that the findings are trustworthy.

Types of evidence

The forms of research used to produce evidence in the human services sector can be divided in several ways. But to begin, we will examine the most fundamental distinction between them: their implicit conception of how knowledge is defined, otherwise known as the *epistemological paradigm* to which they belong (Ryan, 2018). Understanding these base differences between research methods is helpful when determining which one is most appropriate for the type of question asked.

Positivist and interpretivist approaches to research

There are many important questions to ask when evaluating the impacts of a program or service. Some of these questions will relate to a program's effectiveness. For example:

- Has a particular program reduced recurring reports of child maltreatment in participating families?
- Did a particular program increase academic engagement among participants?
- Are children at risk of being placed in out of home care who participated in a particular program more likely to stay in their existing homes than the children involved in an alternative program?

Implicit in these 'effectiveness' questions is a belief that, through rigorous data-gathering and statistical analysis, an objective, observable fact about the world can be uncovered (e.g., one program is, on the whole, more effective at keeping children in their family home than an alternative program). This aligns with the philosophy of positivism, which posits that genuine knowledge can only be derived from empirical data that are collected and analysed as rationally and objectively as possible (Ryan, 2018).

It is also important for service providers to ask questions about how users' experience the program or service. For example:

- What aspect of a particular program do participants consider most helpful/unhelpful?
- Does the target population have specific needs that this particular program does not adequately address?

These sorts of questions (see Figure 1 for more examples) do not attempt to uncover objective facts. They align more closely with the philosophy of interpretivism, which, unlike positivism, emphasises the inherent subjectivity of knowledge and the socially constructed nature of reality (Sale et al., 2002).

Figure 1

Positivist and interpretivist research questions

	Positivism *Objective empirical facts*	**Interpretivism** *Subjective, constructed reality*
Example Research questions	Does this program increase engagement with education? Does participation in this program lead to significant improvement in measurable outcomes for children and families? Are identified outcomes significantly better than if they had not participated in this program, or had participated in a different program?	What do our clients see as their primary support needs? What do the children and families we work with think about this program and what are their perceptions of its impact? What are the lived experiences of children within our service? What aspects of our program do parents and carers consider most/least helpful?

It is important to note that a positivist definition of knowledge should not be considered more correct than an interpretivist definition, or vice versa. However, which is more applicable in a given circumstance depends on the question asked. As a simple heuristic, if you are seeking objective facts about the world (e.g., is program A more effective than program B in delivering certain outcomes in certain circumstances?), then you are asking a question through a positivist lens. But if you are interested in understanding subjective experiences (e.g., how do users of program A feel supported by the included services?), then you are adhering to an interpretivist view of knowledge. We will discuss how these different philosophical approaches to knowledge link to different research methods below.

Quantitative and quantitative research methods

Research methods in the social sciences are typically divided into two categories: quantitative and qualitative. Generally speaking, questions that are posed through a positivist lens are addressed using quantitative research, while questions couched in interpretivism are carried out using qualitative research methods. This is an oversimplification, but a useful starting point.

Quantitative research methods are most suitable for answering positivist-type questions. This is because, through statistical analysis, the numerical data they generate can be used to provide evidence for whether or not there are significant, observable differences between groups (such as between a group that receives treatment and one that doesn't) that are not due to chance or biased sampling.

Figure 2

Experimental designs: some common terms

Intervention group:	This is the study group that participates in the program being evaluated. The difference between an RCT and other experimental designs is that in an RCT, participants are allocated to this group (and to the control or comparison groups) randomly.
Control group:	The control group is made up of participants in a similar situation who do not participate in the program. The control group often receives 'business as usual' services. If participants in the intervention group have better outcomes than those in the control group, this provides evidence that the program is effective.
Comparison group:	If two or more programs are being compared, comparison groups made up of participants enrolled in each of the alternative programs can be used to compare their relative efficacy. In this case, it is still a good idea to have a 'no program' control group so you can assess whether any (or all) of the programs achieve better outcomes than doing nothing at all, or business as usual.
Before and After program comparisons:	This refers to measuring an outcome (such as parenting confidence levels) both before and after program participation to see whether there has been any improvement. However, this type of comparison does not provide strong evidence that the program itself is responsible for any changes in the outcome: parenting confidence levels may have increased simply because more parenting experience was gained over that time.

Quantitative analysis can also examine causality. For example, if you are delivering a parenting program, and you have data that shows that parents improve in their sense of parenting competence over time, you will also want to know whether this improvement was directly because of the program, and not because they would have improved over time anyway as their child grew. To establish causality, researchers use experimental designs that include 'control' or 'comparison' groups (see Figure 2 for a description of some common terms used in experimental research). In the Human Services, the most commonly used experimental designs are Randomised Controlled Trials (RCTs), quasi-experiments, and pre-post designs. While all of these can be used to assess the effectiveness of an intervention (in other words, was the improvement in outcomes due to the intervention itself or other factor?), the level of confidence with which you can establish causality varies between them. See Table 1 for a more comprehensive comparison of research designs.

Quantitative research can also be non-experimental. For example, cross-sectional studies are used to reveal the prevalence of certain health conditions or behaviours (e.g., smoking) across different geographic or socioeconomic zones. This provides very useful data to target interventions to where they are needed most, but it does not provide causal evidence (i.e., why there are more smokers in one suburb than another).

Qualitative research methods are most appropriate when seeking subjective and descriptive accounts of participants' experiences. These includes methods such as focus groups and interviews that aim to capture the experienced 'reality' of their participants. When gathering qualitative data, researchers collect a range of beliefs and opinions held by participants on the topic of interest and synthesise them into a format that attempts to concisely capture the range, depth and complexity of stakeholders' experiences. Qualitative research can be used to develop theories about causality, that may then be tested through quantitative research methods. Qualitative research can be used to inform future research, reveal weaknesses in a program's implementation, and provide directions for future program development (Busetto et al., 2020).

Case study 1: Using an experimental design to assess causality

Are parenting outcomes improved through a volunteer home-visiting program?

TeEACH researchers conducted a study in partnership with Australian NGO's Karitane, The Benevolent Society and Save the Children Australia (Grace et al., 2019) to look at the effectiveness of a volunteer home visiting program for families who were feeling isolated or just in need of extra support. We wanted to know whether being part of this program made a difference to parenting outcomes such as sense of competence and their social connectedness. We decided to use an experimental design (an RCT) with a control group so we could be certain that any positive results we found were because of the program, and not because parent concerns settled over time with experience or as their children got older.

When families were referred to the program, they were asked to be part of the study. If they agreed, we randomly assigned them to receive the program (intervention group), or to not receive the program and continue with the usual supports available within the community (control group). Over a 12-month period, we found that the families who received the program had significantly better parenting outcomes than the control group. And so we can say with confidence that the volunteer home visiting program really does make a significant difference and supports positive outcomes in a way that would not happen otherwise. We can only make such a strong statement with confidence because we had a control group.

Mixed methods research

Mixed method research – research that combines quantitative and qualitative research methods –has gained popularity in Human Services research in recent years as it allows researchers to draw upon the strengths of both qualitative and quantitative research in a complimentary way (Bamberger, 2012). A research project may begin with one-on-one interviews with stakeholders and focus groups in order to gain a more nuanced understanding of the needs of a particular population or of the experiences of program participants. These insights can then be used to inform the design of subsequent quantitative experiments, ensuring that the right questions are asked and that the research is carried out in a culturally sensitive manner. Quantitative research may also come first; for example, quantitative methods may be used to determine that there has been some improvement in a cohort of students, and then qualitative methods are employed to explore the student experience of a program and what they perceive to be the elements of the program that made the most difference. Alternatively, qualitative and quantitative data can be gathered concurrently. While approaches to analysing and integrating quantitative and qualitative data vary considerably (Bazeley, 2009), the overall aim is to develop a more nuanced understanding of the subject of study by not constraining the researcher to posing questions exclusively through a positivist or interpretivist lens (Halcomb & Hickman, 2015).

Case study 2: Qualitative research for program development

Creating a service model for a school wellbeing hub

TeEACH researchers collaborated with NCOSS and a public school in Greater Western Sydney to co-design a service model for a wellbeing hub at the school. Focus groups were held with different groups of stakeholders: children; parents and carers; service professionals (including teachers); and community leaders. Data was collected in adult focus groups through conversations about the strengths of the local area, the service needs, and aspirations for the children's future. The conversations followed the lead and interest of the participants. In the children's focus groups, children were asked to draw a picture or map of their community, and then to identify what they liked, and what was missing. The pictures themselves were research data, and so was the record of the conversation about these pictures as they were being drawn. After the focus groups, a co-design workshop was held. The focus groups data was analysed, and the results used to inform the co-design workshop. Representatives of all stakeholder groups participated in a co-design workshop. The outcome of the workshop was developed into a program logic for the wellbeing hub.

Case study 3: Qualitative research to learn from practitioner expertise, client and community voice

Acknowledgement of Country in early childhood education and care

TeEACH researchers worked with the six early learning centres in the Western Sydney University early learning service (WSUELL) to investigate how they had integrated Acknowledgement of Country into their physical spaces and practices (Grace et al., 2021). Data was collected through individual and group interviews with educators, a focus group with parents and carers, and focus groups with children. The educators were asked about what they did in their centre to Acknowledge Country, and embed Indigenous knowledges in the program. Parents were asked about their understandings and perceptions of these activities. Focus groups with children asked them about artefacts found in their centre, such as Aboriginal and Torres Strait Islander flags, books, and artwork, to generate conversation about what children saw as important and meaningful in learning about Indigenous knowledges and Acknowledgement of Country. Findings from the research were discussed with the Elder Advisory Group at the university. One outcome of the research is the design of online learning materials for educators, to support them and their early learning services to introduce Indigenous knowledges in a reflective, collaborative, child-centred way, based on the principles of respect, responsibility and reciprocity.

Other sources of knowledge

Many of the decisions made within human services – from on-the-fly decisions made in the field by practitioners, to choosing a program to implement – require research evidence to be integrated with forms of knowledge that exist outside of research (Proctor & Rosen, 2008). One such form of knowledge is practice wisdom – practitioners' knowledge, gained via their experience working on the frontline. Practice wisdom is an integral part of effective evidence-based practice, allowing practitioners to tailor their decisions and deployment of evidence-based practices to the individual needs of their service user and their circumstances (Mccracken & Marsh, 2008).

Another vital form of knowledge prominently discussed in human services is the individual experiences of the service user, or client. This is often referred to as client voice or adopting a person-centred approach. A person-centred approach involves elevating, engaging, and valuing the experiences and expertise of the client. The client's voice might be expressed during contact with practitioners or carers, and might be verbal or non-verbal (Department of Health and Human Services Victoria, 2019). In practice, using client voice as a form of knowledge involves abiding by practices and procedures that promote respect and collaboration, ensuring that client perspectives are integrated into decisions (Waters & Buchanan, 2017).

The importance of combining other forms of knowledge with research evidence has been recognised in the scientific literature (Donaldson et al., 2009; Humphries, 2003) and by government bodies. For example, the NSW Department of Communities and Justice evidence portal (2019) explicitly states that decisions should draw on research evidence with practitioner expertise, local knowledge, lived experience, and client voice.


ABOUT

Table 1

Research methods in the Human Services Sector

	Research Method	Strengths	Weaknesses
Qual/Quant	**Systematic reviews and meta-analyses:** These are 'desktop' research approaches in which no new data is gathered. Instead, they examine the strength of the collective evidence about a program or practice approach. While both qualitative and quantitative research can be systematically reviewed, only quantitative research can be included in a meta-analysis. This is because meta-analyses involve estimating intervention effects by collating and analysing the results of multiple studies using statistical methods.	• They help us uncover whether different studies have had different findings regarding a program's effectiveness, and whether the same program has demonstrated the same level of effectiveness in diverse contexts and under diverse conditions. • If a meta-analysis or systematic review shows that a program is effective across multiple sites and contexts, this might suggest that the program is suitable for implementation in diverse contexts.	• Systematic reviews cannot tell you about the effectiveness of programs as they are currently implemented. For example, if a parenting program like Triple P is being delivered, we will know from this approach that Triple P has a strong evidence base. However, we won't know if our implementation of Triple P is having the same impact on our participants–we can just be confident that it has been impactful for other families in previous research.
Quant	**Randomised Controlled Trial (RCT)** RCTs are a rigorous, gold standard experimental research design that examines the effect of a program on a cohort of clients in relation to a control group. To ensure that any observed differences are due to the treatment or intervention and not to other factors, participants are randomly assigned to either a treatment or control group. A program is considered effective if, on average, program participant outcomes are statistically significantly better than of those in the control group.	• RCTs provide the strongest evidence in response to questions such as "is this program actually effective at bringing about the change we want?". When well designed, they minimise the effects of biases and confounds that can skew results in misleading directions. • If the research methods have been described with adequate depth and transparency, published RCTs can be carefully scrutinised, leading to greater confidence in their findings.	• The results that are produced are based on specific implementation conditions and contexts, and it can be difficult to assess the extent to which the same results would be produced in a different context. • Tend to require large participant numbers. • RCTs are more financially and logistically demanding than other research methods. • Some human service professionals feel uncomfortable with the idea that some clients will be randomly assigned to a 'control' group – meaning that they won't receive the program (at least initially), and so it can be difficult to secure the support of practitioners in the conduct of an RCT.
Quant	**Quasi-Experiment** An experimental research design that differs from RCTs by not having participants randomly assigned to the intervention or control conditions. Instead, a comparison group is established, made up of people who were not looking to receive the program anyway. Quasi-experiments are generally used when RCTs are deemed inappropriate due to logistical, ethical or resource constraints.	• Quasi-experiments can be simpler and more affordable to plan and execute than RCTs. However, conducting a well-controlled quasi-experimental is still a challenging undertaking. • By including a comparison group, quasi-experiments allow researchers to rule out explanations for improved outcomes such as the passing of time, child maturation, etc.	• The fact that one group of participants (the intervention group) chose to participate in the program while the other (the control group) means that, inevitably, there is a fundamental and potentially confounding difference between the groups. • Without random allocation to groups, balancing potential confounding factors such as location, socioeconomic status, age, cultural background, etc. across the groups can be challenging and may contaminate the results.
Quant	**Single group pre-post design** An experimental design in which program effectiveness is assessed by comparing outcome metrics taken prior to program participation with those taken during or after completion. Control or comparison groups are typically not included because individuals become their own control	• It is the simplest and most cost-effective experiment design for assessing program effectiveness. • Requires fewer participants than experimental designs with multiple groups.	• Causal inferences cannot be made with a high degree of confidence because you cannot rule out the changes in outcomes were due simply to the passage of time, and not the program itself. This is of particular concern when researching young people because we know that child and adolescent development can be quite rapid and that maturation effects can influence outcomes.

Qual/Quant	**Observational studies** This is a catch-all term for various quantitative and/or qualitative non-experimental research designs (e.g., cohort studies, case-controlled studies, and cross-sectional studies) in which researchers gather their data without intervening in the program and without allocating participants to treatment conditions. Data may be collected retrospectively from records, via field observation, or through more direct forms of enquiry such as questionnaires.	• Are ideal for evaluating implementation fidelity, which is the degree to which a program is implemented as intended. • Are useful for providing a descriptive account of complex environments and circumstances for the purposes of exploration or for guiding future research.	• The lack of researcher control over participant allocation limits the confidence by which causal claims can be made.
Qual	**Action research** Refers to systematic, practitioner-led research carried out for the purposes of deepening knowledge, evaluating practices, and for driving future improvements. It requires practitioners to develop a range of research and analysis skills, including critical self-reflection.	• Empowers practitioners to lead the development of their practice, directly benefiting both their clients and themselves. • Encourages a systematic and rigorous approach towards understanding and addressing problems in practice.	• The additional workload can be challenging for practitioners to manage. • Can lack the rigor of other research methods.
Qual	**Case Study** Through in-depth examinations of participants (it may be one participant, or multiple), case studies aim to generate nuanced and detailed accounts of their subjects. This can help you better the experiences of program participants as well as what forces shape this experience.	• Useful for developing a deeper understanding of the subjective experiences of participants and the complex socio-environmental systems in which they exist. • The level of detail of a case study makes them useful as teaching aids and for communicating the experiences of participants.	• Are unsuitable for making causal claims regarding program efficacy, although analysis of data can develop theories about how programs create change. • The outcomes and experiences may not be generalisable beyond the case study site or case study participants.
Qual	**Interviews** Are used to collect detailed accounts of the subjective experiences and beliefs of participants in a systematic manner. Interviews can be open-ended and participant led, structured around specific questions, or lie somewhere in between.	• The one-on-one intimacy of an interview can allow the interviewer to build a comfortable rapport with the subject. This may yield more valuable insights than in a focus group setting.	• The time and labour required to conduct and analyse large numbers of interviews can be prohibitive. • The subjective experiences of interviewees cannot be considered to be representative of a larger population, or generalised across different populations.
Qual	**Focus Groups** Typically involve up to eight participants in a structured discussion around a specific set of questions. Discussion is moderated by the facilitator/researcher.	• The exchange of ideas between diverse participants can yield unique insights and may be more engaging than interviews for some participants. • They are more efficient at generating detailed insights from large numbers of participants than interviews.	• Ensuring that speaking time is not distributed disproportionately can be challenging. • Participants may be vulnerable to 'groupthink'.

EVALUATING EVIDENCE

What does it take for a program to be declared 'evidence-based'?

The expectation that experimental research methods (the gold standard being RCTs) should be used to evaluate human services is higher than ever before (Baron, 2018; Gray et al., 2009). Using program effectiveness as their primary evaluation metric, program clearinghouses, such as the California Evidence-Based Clearinghouse for Child Welfare and the Early Intervention Foundation, review swathes of programs and assign evidence ratings. If a clearinghouse finds sufficiently rigorous experimental evidence demonstrating that a program improves (at least some) outcomes, then it will receive the clearinghouses' stamp of approval as an EBP (see Table 2).

Table 2

An example of an Evidence Rating Scale developed by the Centre for Evidence and Implementation (2017) on behalf of the Department of Health and Human Services Victoria.

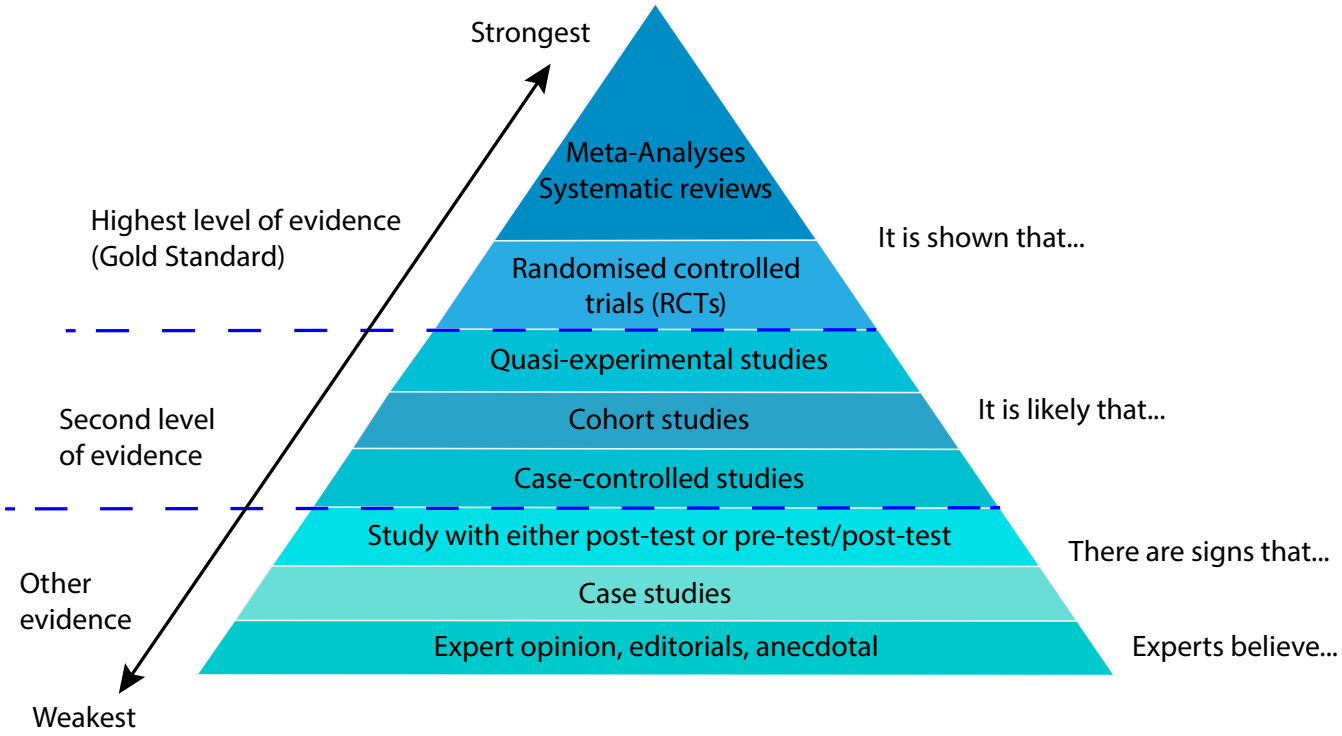
Rating	Description
Well supported by research evidence	At least one high quality **systematic review** that statistically synthesises the outcomes (i.e., a systematic review that is inclusive of a meta-analysis) based on RCTs and/or quasi-experimental studies, and conducted in a range of usual care or practice settings, has found the practice or program to be **superior** to an appropriate comparison practice. In this systematic review, the practice or program has been shown to have a **sustained effect** for at least one primary outcome lasting at least one year beyond the end of treatment, when compared to a control group. **No adverse effects** were identified for the practice or program.
Supported by research evidence	At least **two rigorous RCTs** conducted in a usual care or practice setting has found the practice or program to be **superior** to an appropriate comparison practice or program. In both studies, the practice or program has been shown to have a **sustained effect** lasting at least one year beyond the end of treatment, when compared to a control group. **No adverse effects** were identified for the practice or program.
Promising research evidence	At least one **rigorous RCT or quasi-experimental study** conducted in a usual care or practice setting has established the program's or practice's benefit over the control group or found it to be **superior** to an appropriate comparison practice. **No adverse effects** were identified for the practice or program.
Emerging research evidence	The practice or program meets the requirements of the logic-informed rating. The practice or program has been tested using high-quality quasi-experimental or other study designs. Testing of impact is underway but evidence needed to meet the requirements of the 'promising research evidence' rating is not yet achieved.
Logic-informed	Key elements of the logic model / program logic for the practice or program are defined and verified in relation to the practice or program and the underpinning scientific evidence (adapted from Early Intervention Foundation, 2016).
Evidence fails to demonstrate effect	Two or more randomised, controlled outcome studies have found that the practice or program **has not resulted in improved outcomes**, when compared to usual care. If multiple outcome studies have been conducted, the overall weight of evidence does not support the benefit of the practice or program
Concerning practice	If multiple outcome studies have been conducted and one or more of the following applies: • The overall weight of evidence suggests the practice or program has a **negative effect** upon clients served • There is **case data** suggesting a risk of harm that: (a) was probably caused by the practice or program; and (b) the harm was severe and/or frequent • There is a **legal or empirical basis** suggesting that, compared to its likely benefits, the practice or program constitutes a risk of harm to those receiving it.

Evidence hierarchies

Evidence hierarchies (see Figure 3) are frameworks for ranking the relative strength of findings obtained from different research methodologies. Originally devised as a means by which to assess evidence for treatment efficacy in medicine, research methods are ranked according to the strength by which causal claims regarding effectiveness (i.e., treatment caused X to improve) are supported (Ho et al., 2008).

Figure 3

An example of an evidence hierarchy sourced from the NSW Department of Communities and Justice (2020).



Evaluating a program in accordance with an evidence hierarchy has some clear advantages. For program clearinghouses or other evaluators, adhering to a simple and standardised evidence ranking system enables the voluminous literature on human services to be examined and evaluated systematically and efficiently. This is not trivial – the faster programs can be assessed and the findings disseminated, the faster the critical transition from "knowledge production to knowledge utilisation" can occur (Soydan et al., 2010, p. 1).

However, critics have noted the limitations to adhering strictly to conventional evidence hierarchies. Evidence hierarchies restrict the scope of what is considered legitimate evidence. Firstly, according to Gray and colleagues (2009), the 'evidence-based' movement has inadvertently entrenched unnecessarily restrictive and simplistic notions of what is good and bad evidence. The evidence hierarchy, which – at face value – is simply a ranking of research methodologies based on their internal validity, has become synonymous with evidence, more broadly. In the process, client experience and practitioners' practice wisdom are undervalued. Evidence hierarchies also fail to capture the synergistic relationship between methods high on the evidence pyramid (i.e., RCTs) and other forms of evidence. RCTs are often informed by preliminary research, and successful program implementation, supported by RCTs, requires evidence gathering methods outside of those endorsed by evidence hierarchies.

Secondly, evidence hierarchies are important to inform program selection and implementation, but not in the absence of information about context. The designation of a program as 'well supported by research evidence' does not equate to a recommendation that the program is appropriate for adoption and/or adaptation. The information required for successful implementation, such as its generalisability, flexibility, and appropriateness for the target population are not a focus of evidence hierarchy-based evaluations (Nutley et al., 2013). For example, a program developed with urban families in a southern state of the United States of America might not be appropriate for Indigenous families in rural parts of Australia, even if it does qualify as 'evidence-based'. To make an informed decision, much more information about the extent to which the program can be adapted to meet local needs is required.

Case study 4: When considering an EBP, the details matter

In 2006, Gardner and colleagues published a paper reporting on a rigorous randomised controlled trial to test the effectiveness of a parenting intervention program (The Incredible Years) in a community setting, delivered by trained volunteers, to support parents of children with conduct problems. The program was designed to:

- Reduce the child's problem behaviours
- Increase positive parenting behaviours
- Reduce negative parenting behaviours
- Increase parenting skills
- Increase parenting confidence
- Decrease parent depression, and
- Reduce problem behaviours in siblings

In their study, they found that there were significant results for each of the above outcomes, except for the outcome relating to parent depression. And so, if a service provider is looking for a program that will address parental depression, this is not the program for them (or at the very least it will need to be supplemented).

The term "Evidence-based Program' is often used as a blanket statement or rubber stamp to endorse a program in full. However, it is always important to look at which outcomes actually were supported by the research evidence, and to ensure that these outcomes align with the objectives and needs of local services.

An evidence matrix

While an evidence hierarchy is useful to rank research methods by how rigorously they control bias, it is less useful when applied to questions beyond effectiveness, such as whether a program will fit the target population, why a given program does (or does not) work, or what facets of a program can be adapted to meet the needs of a local context without compromising program fidelity.

A more useful means to conceptualise the applicability of different types of evidence is presented in an evidence matrix (see Figure 4). By providing evidence rankings that differ depending on the question, the evidence matrix suggests that different research methods are not innately superior or inferior to others – it all depends on the question. Petticrew & Roberts (2003) called this approach 'methodological pluralism' and claimed that it better represents the relationship between the questions being asked and the research method most suited to answer it. Furthermore, it resists privileging questions of efficacy above all the other sorts of questions that one may ask about a program.

Unsurprisingly, for questions regarding effectiveness (e.g., does doing this work better than doing that?) the evidence rankings mirror the evidence hierarchy: systematic reviews provide the most robust form of evidence, followed by RCTs and quasi-experimental designs. However, for questions relating to a program's mechanisms of action, client experiences, or its overarching purpose and meaning, the evidence hierarchy is inverted in favour of methods that serve to collect qualitative data.

Figure 4
An evidence matrix (sourced from Nutley et al., 2013). Adapted from Petticrew and Roberts (2003).

Research Question	Qualitative research	Survey	Case-control students	Cohort studies	RCTs	Quasi-experimental studies	Non-experimental studies	Systematic reviews
Does doing this work better than doing that?				+	++	+		+++
How does it work?	++	+					+	+++
Does it matter?	++	++						+++
Will it do more good than harm?	+		+	+	++	+	+	+++
Will service users be willing to or want to take up the service offered?	++	+			+	+	+	+++
Is it worth buying this service?		++						++
Is it the right service for these people?	++	++	+	+				+
Are users, providers, and other stakeholders satisfied with the service?	++							

USING EVIDENCE

Reducing the gap between research and practice is the central focus of implementation science: 'the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practice into routine practice, and hence, to improve the quality and effectiveness of health services and care' (Eccles & Mittman, 2006, p. 1). We begin with an overview of an implementation science innovation known as The Heptagon Tool. This tool is designed to help you ask the right questions about current or prospective programs, ensuring that you are aware of what evidence is required to make well-informed decisions.

Following The Heptagon Tool, we provide an overview of common elements – a flexible approach to incorporate research findings into program development. The Heptagon Tool and common elements approach are complementary: while The Heptagon Tool prompts you to gather evidence to answer important questions, the common elements approach broadens the pool of evidence to draw from.

To finish, we discuss the benefits of embedding ongoing research into a program's wider implementation strategy.

The Heptagon Tool

Asking the right questions

The failure of well-intentioned programs–whether due to limited uptake among the target population, financial or logistical pressures, or simply because no tangible benefits were produced–can often be traced back to decisions around implementation that were not sufficiently well researched or informed by evidence. At face value, being aware of a community need and deciding to do something to address it seems straightforward; however, understanding the needs of communities requires careful thought and research. Specifically, it requires us to ask the right questions (about the target population, their community, the proposed intervention, and service capacity) and seek answers in a robust and ethical manner. The Heptagon Tool (see Figure 5) can help you to ask a range of useful questions about a program's fit and feasibility.

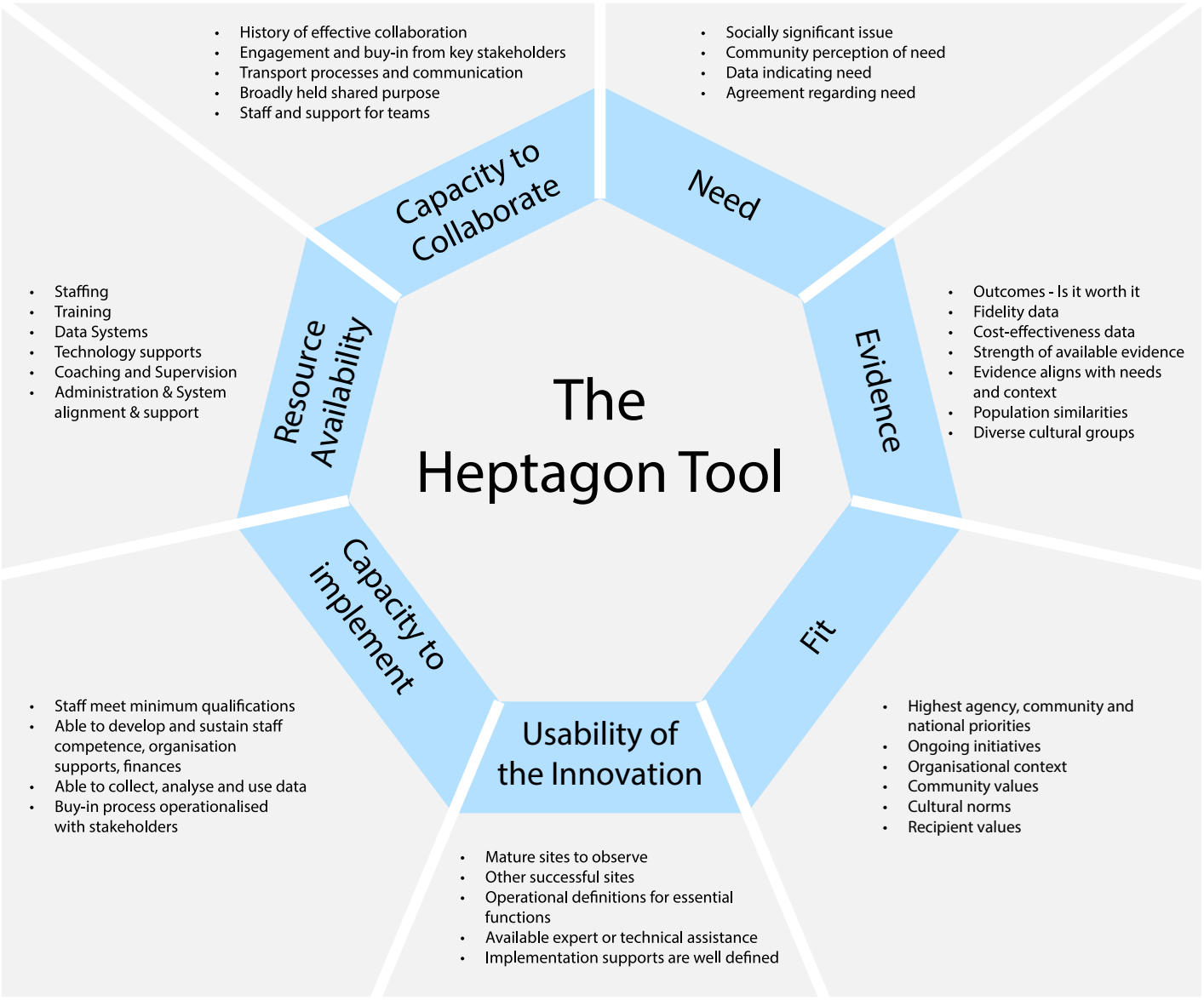
How it works

The Heptagon Tool consists of seven categories of enquiry, referred to as *dimensions* (see figure X). The dimensions **need**, **evidence**, and **fit** are concerned with whether a program *is the right thing to do* – is it appropriate for the needs of the target population? Is there evidence that this program effectively addresses this need? Is it culturally appropriate?

The dimensions ***capacity to implement, capacity to collaborate, resource availability*** and ***usability of the innovation*** are concerned with whether a program *can be done in the right way* – do we have the capacity and resources required to implement the program? Is the decision-making around adopting this program included all relevant stakeholders?

The Heptagon Tool has been designed to ensure that questions about a program's efficacy and appropriateness (i.e., the '*is this the right thing to do?'* questions) are balanced with questions that ask whether the program can be successfully implemented (i.e., the '*can it be done in the right way?*' questions). This is a critical point: there is no point adopting an evidence-based program if you lack the resources required to implement it with fidelity, or if you are unable to get the relevant stakeholders to support the initiative.

Figure 5
An evidence matrix (sourced from Nutley et al., 2013).



To see the full range of questions posed by The Heptagon Tool, see the supporting documentation (Van Dyke et al., 2019). See Table 3 for a concise sample of these questions, as well as the appropriate forms of information (evidence-based or otherwise) that are necessary to answer them.

Table 3
Sample questions from The Heptagon Tool (Van Dyke et al. 2019)

		Sample Question	Suggested evidence
Is it the right thing to do?	Need	*What data demonstrate the need for an innovation or improvement? What does the data tell us about the underlying needs of the population to be served?*	-**systematic reviews** -**Qualitative** (i.e., on-site surveys, focus groups, interviews) -**Quantitative** (i.e. cross-population comparison data)
	Evidence	*What is the strength of the available evidence, and does it align with the needs and context of the target population?*	-**systematic reviews** -**Quantitative** (evidence of effectiveness from RCTs and quasi-experiments). -**Qualitative** (i.e., on-site surveys, focus groups, interviews)
	Fit	*To what extent does this innovation or improvement effort align with one of the highest agency, community, or national priorities?*	-**Qualitative** (i.e., on-site surveys, focus groups, interviews) -**Grey literature** (e.g., Government policy papers, funding guidelines)
Can we do it in the right way?	Capacity to collaborate	*Have key stakeholders been engaged in the exploration of this innovation or improvement effort?*	-**Qualitative** (i.e., on-site surveys, focus groups, interviews)
	Resource availability	*Are the necessary resources and supports available to meet the programmatic requirements for the innovation or improvement effort?*	Requires an analysis of program requirements and organisational capacity, not evidence
	Capacity to implement	*Will it be possible to sustain the necessary support to maintain staff competence, to maintain organisation supports, to financially maintain this innovation or improvement effort?*	Requires an analysis of program requirements and organisational capacity, not evidence
	Usability of the innovation	*Is expert or technical assistance available?*	No evidence required

When to use it
The Heptagon Tool can be used at various points in a program's lifecycle. In the exploration phase (in which community needs are identified, prospective programs are compared, and feasibility is assessed) The Heptagon Tool provides clear criteria to evaluate different program options; this can help to structure and clarify the decision-making process.

The Heptagon Tool can also be used to implement or reassess a program. In these cases, it acts more as a diagnostic tool to determine whether the program and local needs remained aligned, or whether it needs to be adapted or ceased to reflect local changes in need and context.

Common elements

What are common elements?

Common elements are practices and techniques that are common across multiple experimentally verified programs and interventions. These practices are the active ingredients of EBPs that, once isolated and defined, can be delivered outside of the program structures from which they were drawn, either individually or in combination (Barth et al., 2012).

This modular and flexible approach to program development is useful in circumstances when implementing an off-the-shelf EBP is not possible or appropriate – for instance, due to resource limitations, cultural misalignment between the program's context of origin and its intended implementation site, or other factors. The common elements approach now forms part of the broader strategy adopted by several state governments, including in NSW (NSW Department of Communities and Justice, 2022b) to increase the degree to which human services base their practices on research evidence.

To identify common elements, relevant literature is compiled through a review process and then analysed to identify and extract the common elements (Chorpita et al., 2005). For an example set of common elements, see Table 4.

Case study 5: identifying common elements for preventing child maltreatment

The NSW Department of Communities and Justice (DCJ) recently carried out a project that aimed to identify common elements for preventing child maltreatment (2022c).

The project began with the formulation of the following question:

"What programs are effective in preventing child maltreatment in families?"

Informed by this question, the following was carried out:
An evidence review was conducted, and a list of EBPs shown to effectively improve outcomes relating to preventing child maltreatment was compiled
The core practices within these program were extracted using a content analysis approach
The practices most commonly shared between the programs were identified as 'common elements'

This project produced a repository of common elements for preventing child maltreatment (subdivided into 'core components' and 'flexible activities') that includes practical implementation guides for each common elements (NSW Department of Communities and Justice, 2022c). See the DCJ Evidence Portal Technical Specifications for more details on the research process (NSW Department of Communities and Justice, 2022d).

Table 4
10 Common elements for working with young people and families (SA Department of Human Services, 2022b)

Group	Common Element
Building Engagement	Open-ended questions, Affirming, Reflective listening, and Summarizing *(OARS)*
	Checking for Understanding
	Seeking Feedback
Preparing for change	Collaborative Partnership
	Building Motivation for Change
	Identifying Priorities
	Goal Setting
Enhancing Family Functioning and Safety	Communication Skills
	Increasing Safety Through Conversations
	Building Capacity for Family Led Problem Solving
Practice guides for each of the common elements can be found on the SA Govt Human Services Early Intervention Research Directorate page (SA Department of Human Services 2022b)	

Strengths and limitations of the common elements approach

Strengths:

- The ability to flexibly and dynamically combine and implement common elements is well suited to the complex and shifting circumstances in which human services are often delivered. (Barth et al., 2012)
- The time taken to prepare and deploy services constructed around common elements is likely to be faster than the time taken to adopt and implement an EBP. (NSW Department of Communities and Justice, 2022b)
- Practitioners trained in common elements develop core capabilities and skills that can be used across different programs. This helps to standardise practices across the service system and results in more efficient training procedures. (Chorpita et al., 2005)
- There is some evidence suggesting that practitioners prefer delivering modular programs over EBPs. (Borntrager et al., 2009)

Limitations:

- While there is some evidence suggesting that programs constructed from common elements can produce sustained positive outcomes (Chorpita et al., 2013) and faster improvements (Chorpita et al., 2017) than manualised program controls, the common elements themselves have not been evaluated independently of the EBPs from which they were extracted. Accordingly, there is no research evidence supporting their effectiveness either in isolation or in novel combinations, despite claims that common elements are 'grounded in evidence'. (Barth et al., 2012)
- Just like the EBPs from which they were derived, the strength of evidence underlying common elements is compromised when they are drawn from EBPs that have only received evaluations in contexts or with populations that differ significantly from where they are to be implemented. (SA Department of Human Services, 2022b)
- Unlike with EBPs, a program logic, theory of change, and implementation frameworks might need to be developed from scratch. This can be a considerable undertaking, but is critical if common elements are to be implemented in a thought-out, targeted and systematic manner. (Barth et al., 2012; Knight et al., 2018)

How are common elements used?

Program creation. When designing a new program, common elements can be used in a modular fashion to flexibly shape program practices to fit local needs. For programs built in this fashion, Knight and colleagues (2018) emphasised the need to develop a clear program logic to guide the selection of common elements in accordance with strategic aims – this ensures the elements are compatible in terms of implementation.

Program evaluation and development. Common elements offer a novel method to introduce evidence-informed practices into pre-existing program structures and assess whether program practices are aligned with best practice in the research literature. This is done by deconstructing a program into its constituent techniques and practices and comparing these to common elements identified in the literature. For some more examples, explore the resources listed in Table 5.

Table 5

Common elements resources

Topical common elements reviews	Preventing Child Maltreatment	NSW Department of Communities and Justice, (2022c)
	Child Maltreatment interventions	Van der put et al. 2018
	Placement Prevention	Lee et al. 2014
	Out-of-school academic interventions for at-risk children	Engell et al. 2020
	Youth Mentoring	NSW Department of Communities and Justice. (2022a)
	Child and Family Support System	SA Department of Human Services. (2022a)
Databases of common elements	Database of common elements of mental health-based treatments for youth	PracticeWise (2023)
	Database of common elements of mental health treatments	CETA (2021)
Common elements-based Programs	A modular program for treating youth mental health in community settings	Chorpita & Weiss, 2012

Building evidence through continuous evaluation

To determine whether a program comprised of common elements produces the desired outcomes, it is imperative to evaluate the program, ideally in an ongoing fashion, during implementation and beyond. While incorporating ongoing evaluation procedures is important when delivering programs constructed around novel and unproven combinations of common elements, the same applies when administering an EBP. It is impossible to guarantee that an EBP will replicate its success in your context, even if it is implemented with high fidelity. We discuss the concept of ongoing evaluation in more depth in the following section.

Embedded research

There is widespread acknowledgement that working effectively within complex and dynamic social, cultural, and environmental systems requires us to be nimble – we must continually learn, adapt, and respond to changing circumstances and priorities (Churruca et al., 2019). It is not surprising, then, that there are calls to shift away from sole reliance on one-off external evaluations as the primary source of evidence, towards continuous research, embedded into a program's broader implementation strategies, providing ongoing feedback to drive program development (Barry et al., 2018; Churruca et al., 2019; Vindrola-Padros et al., 2017).

Embedded research is typically implemented through partnerships between service providers and research institutions and/or through the employment of onsite researchers. The rationale for embedded research is that research is not a moment in time. To ensure a program remains effective and aligned to current objectives requires us to ask the right questions, keep gathering and assessing data, make iterative adjustments to a program's implementation based on this data, and repeat.

Figure 6
A framework for embedded research (sourced from Barry et al., 2018).

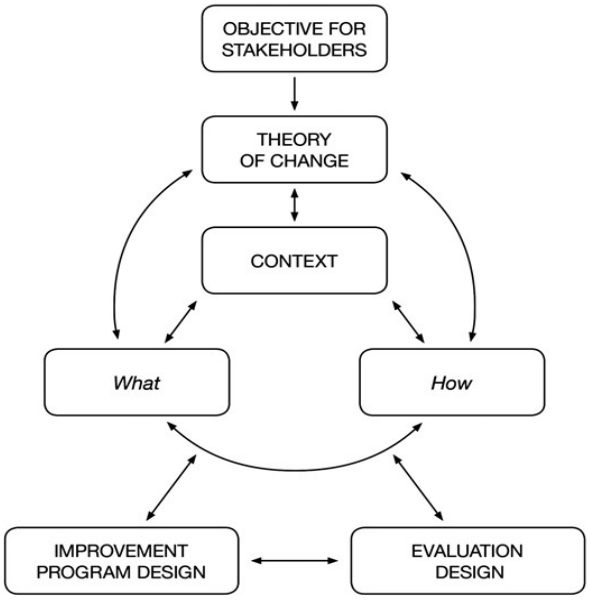


Figure 6 offers a big picture view of how evidence production through research can be embedded into a program's wider structure. The circle in the centre comprising context, theory of change, the what (what is the program doing?), and the how (how is it doing it?), encompasses the program's rationale and implementation practices. Evaluation design (i.e., the embedded research process) sits outside of this circle, but responds to it: it adapts according to the current state of the program while feeding insights back into the program through design improvements.

Advantages for researchers:
Embedded researchers can develop a more sophisticated and intimate understanding of the nuances of program implementation. This can lead to better questions, more buy-in from participants, and research designs that are better adapted to the context (Churruca et al., 2019).

Advantages for implementers:
Embedded research offers many benefits to implementers. Fundamentally, the only way to know whether a program is effective is by continuing to gather evidence. And if a program is not achieving its outcomes, research insights can guide its development. Furthermore, when embedded research is carried out in collaboration with practitioners and other non-research trained staff, it provides them with the opportunity to develop research literacy and skills that are transferable to future programs or positions (Churruca et al., 2019).

Investing in embedded research is often thought of as a luxury for when there are excess funds available. But to achieve the greatest possible positive impact with finite resources, decision-making needs to be guided by evidence.

Case study 6: Using the heptagon tool to work well with evidence

Let's imagine that we are working in a rural community with high levels of socioeconomic disadvantage, where many children are starting school without some of the basic skills and knowledge we would expect them to have. For example, many children are starting kindergarten without being able to count to 10 or without any knowledge of the alphabet. Some of the children don't seem to have been exposed to story books before, or don't know which end of a pencil to use. There are even a small number of children who are struggling with basic skills like toileting. The local AEDC (Australian Early Development Census) data, combined with the observations of teachers, confirms our concern for these children and the need to support their preparation and transitions to formal education. We also know from research that early education experiences can set the trajectory for children's future learning.

We start by looking to understand the **NEED.** We examine the available data, and we talk to local community members, teachers, and parents about their perceptions of the challenges and what they think might be needed to better prepare young children for school. It becomes clear that the children who are struggling most are those who are the children of miners. We learn that these children move with their families frequently and, as a result, generally have not attended early childhood education settings or spent a lot of time with other children. We decide to prioritise working with these children as a first step.

Next, we review the **EVIDENCE.** Researchers and practitioners work together to review which transition to school programs have the strongest evidence base. We search for transition to school programs that are home-based, for children who don't attend early childhood settings, and we look to see whether any have been tested with rural and transient populations. We learn a lot in this review about the best ways to support children and families, and the most important skills for children to develop before they go to school. We decide that the HIPPY program is an EBP that we would like to implement in our context.

HIPPY seems to be a good **FIT** in terms of community values. However, we are worried about whether we can implement it properly. HIPPY is generally delivered over 2 years. We know many of our parents are not in our community for a full 2 years before their child starts school. We decide we'd like to adapt the program so it can be delivered over a shorter period of time. We would also like to incorporate some of our learnings from the literature about directly addressing issues of transience and the impact on child development (drawing on a **COMMON ELEMENTS** approach). We want to feel confident that our adaptations will not compromise program effectiveness, and so we work with researchers in our adaptation and decide to conduct a pilot study, then embed ongoing program evaluation in our delivery.

The **USABILITY OF THE INTERVENTION** seems to be clear, and we quickly are across the elements of the program. We feel confident that it can be delivered within our community, although we do have some concerns about our current **CAPACITY TO IMPLEMENT** because we do not have any workers trained as HIPPY facilitators. We examine our **RESOURCE AVAILABILTY** to ensure we can afford the program, training and support costs.

In all of this work, we have partnered with the local primary school to ensure that we are addressing their goals for the children, and also to incorporate into our program early engagement with the school. We have also partnered with the mining company who have agreed to provide some funding to support the roll-out of the program, and to support the participation of their workers by offering flexible work arrangements so they can attend HIPPY sessions. A partnership has also been established with researchers to ensure rigorous evaluation, review of the literature, and dissemination of our findings so we can support other towns with similar challenges. We have formed a strong consortium and a **CAPACITY TO COLLABORATE.**

CONCLUSIONS

What qualifies as 'good evidence' depends on the question

The term 'evidence-based' has become ubiquitous in the human services sector – it is attached to more programs, practices, and policies than ever before. But when viewed in isolation, it doesn't tell us much at all. "Evidence for what?" would be the logical first question to ask. Program clearinghouses generally won't provide the 'evidence-based' tick of approval unless a program has been subjective to rigorous experimental evaluation and was shown to improve at least *some* participant outcomes. But what were the outcomes that improved, and what ones didn't? And for whom? And where? Answering these questions requires you to examine the research closely, and this takes time.

Although 'evidence-based' is generally understood as being synonymous with evidence for effectiveness, there are many important questions that do not address effectiveness but are nonetheless critical to raise the prospects of a program's overall success and effectiveness. These questions (e.g., how do the children we work with perceive this program? Do they think it is helpful? And if not, why?), typically of a qualitative nature, should be posed at all stages of a project's lifecycle. It is through rigorous qualitative research that you develop your understanding of the context you are working in, the needs of the community from their own perspective, as well as the social and cultural factors that need to be considered.

Which questions are asked matters

Knowing what questions to ask is not as simple as one might suppose. Just as empirical research requires clear and systematic procedures, formulating the right questions to ask should be done through a structured and methodical process. The is where The Heptagon Tool comes in. When used as a guide for your analysis of both the implementation site and your current (or prospective) programs, it can help to ensure that you gather all the information and evidence necessary you need to make well-informed decisions.

Research is not a moment in time

The usefulness of the evidence is determined not only by its quality, but also by its relevance to the present circumstances. True adherence to an evidence-based philosophy involves a continuous cycle of research and improvement, achieved by integrating ongoing research into a program implementation plan. When research and practice are integrated synergistically from the ground up, both benefit: research borne of intimate knowledge is more targeted and relevant, and practice that is responsive to new research insights can continue to evolve and develop in a manner that is guided by evidence.

REFERENCES

Baron, J. (2018). *A brief history of evidence-based policy.* The annals of the American Academy of Political and Social Science, 678(1), 40–50. https://doi.org/10.1177/0002716218763128

Barry, D., Kimble, L. E., Nambiar, B., Parry, G., Jha, A., Chattu, V. K., Massoud, M. R., & Goldmann, D. (2018). *A framework for learning about improvement: Embedded implementation and evaluation design to optimize learning.* International Journal for Quality in Health Care, 30(suppl_1), 10–14. https://doi.org/10.1093/intqhc/mzy008

Barth, R. P., Lee, B. R., Lindsey, M. A., Collins, K. S., Strieder, F., Chorpita, B. F., Becker, K. D., & Sparks, J. A. (2012). *Evidence-Based Practice at a crossroads: The timely emergence of common elements and common factors.* Research on Social Work Practice, 22(1), 108–119. https://doi.org/10.1177/1049731511408440

Bazeley, P. (2009). *Analysing mixed methods data.* In S. Andrew & E. J. Halcomb (Eds.), Mixed Methods Research for Nursing and the Health Sciences (pp. 84-118). https://doi.org/10.1002/9781444316490.ch6

Borntrager, C. F., Chorpita, B. F., Higa-McMillan, C., & Weisz, J. R. (2009). *Provider attitudes toward evidence-based practices: Are the concerns with the evidence or with the manuals?* Psychiatric Services (Washington, D.C.), 60(5), 677–681. https://doi.org/10.1176/ps.2009.60.5.677

Busetto, L., Wick, W., & Gumbinger, C. (2020). *How to use and assess qualitative research methods.* Neurological Research and Practice, 2(1), 14. https://doi.org/10.1186/s42466-020-00059-z

CETA (2021) *Common Elements treatment approach* https://www.cetaglobal.org/

Chorpita, B. F., Daleiden, E. L., Park, A. L., Ward, A. M., Levy, M. C., Cromley, T., Chiu, A. W., Letamendi, A. M., Tsai, K. H., & Krull, J. L. (2017). *Child STEPs in California: A cluster randomized effectiveness trial comparing modular treatment with community implemented treatment for youth with anxiety, depression, conduct problems, or traumatic stress.* Journal of Consulting and Clinical Psychology, 85, 13–25. https://doi.org/10.1037/ccp0000133

Chorpita, B. F., Daleiden, E. L., & Weisz, J. R. (2005). *Identifying and Selecting the Common Elements of Evidence Based Interventions: A Distillation and Matching Model.* Mental Health Services Research, 7(1), 5–20. https://doi.org/10.1007/s11020-005-1962-6

Chorpita, B. F., Weisz, J. R., Daleiden, E. L., Schoenwald, S. K., Palinkas, L. A., Miranda, J., Higa-McMillan, C. K., Nakamura, B. J., Austin, A. A., Borntrager, C. F., Ward, A., Wells, K. C., Gibbons, R. D., & Research Network on Youth Mental Health. (2013). *Long-term outcomes for the Child STEPs randomized effectiveness trial: A comparison of modular and standard treatment designs with usual care.* Journal of Consulting and Clinical Psychology, 81(6), 999–1009. https://doi.org/10.1037/a0034200

Chorpita, B. & Weiss, J. (2012) *Modular Approach to Therapy for Children with Anxiety, Depression, Trauma, or Conduct Problems* https://www.practicewise.com/portals/0/MATCH_public/index.html

Churruca, K., Ludlow, K., Taylor, N., Long, J. C., Best, S., & Braithwaite, J. (2019). *The time has come: Embedded implementation research for health care improvement.* Journal of Evaluation in Clinical Practice, 25(3), 373–380. https://doi.org/10.1111/jep.13100

Department of Health and Human Services Victoria. (2019). *Client voice framework for community services.* https://www.dhhs.vic.gov.au/sites/default/files/documents/201910/client-voice-framework-for-community-services.pdf

Donaldson, S. I., Christie, C. A., & Mark, M. M. (2009). *What Counts as Credible Evidence in Applied Research and Evaluation Practice?* SAGE.

Eccles, M. P., & Mittman, B. S. (2006). *Welcome to Implementation Science.* Implementation Science, 1(1), 1. https://doi.org/10.1186/1748-5908-1-1\

Engell, Kirkøen, B., Hammerstrøm, K. T., Kornør, H., Ludvigsen, K. H., & Hagen, K. A. (2020). *Common Elements of Practice, Process and Implementation in Out-of-School-Time Academic Interventions for At-risk Children: a Systematic Review.* Prevention Science, 21(4), 545–556. https://doi.org/10.1007/s11121-020-01091-w

Grace, R., Woodrow, C., Townley, C., Baker, E., Kaplun, C., & Staples, K. (2021). *Acknowledgement of Country practices in Early Learning Centres.* https://doi.org/10.26183/fw33-zg22

Gray, M., Plath, D., & Webb, S. (2009). *Evidence-based Social Work* (0 ed.). Routledge. https://doi.org/10.4324/9780203876626

Halcomb, E. J., & Hickman, L. (n.d.). Mixed methods research. Nursing Standard: promoting excellence in nursing care, 29 (32), 41-47. http://dx.doi.org/10.7748/ns.29.32.41.e8858

Ho, P. M., Peterson, P. N., & Masoudi, F. A. (2008). *Evaluating the evidence.* Circulation, 118(16), 1675–1684. https://doi.org/10.1161/CIRCULATIONAHA.107.721357

Humphries, B. (2003). *What else counts as evidence in evidence-based social work?* Social Work Education, 22(1), 81–91. https://doi.org/10.1080/02615470309130

Knight, A., Maple, M., Shakeshaft, A., Shakehsaft, B., & Pearce, T. (2018). *Improving the evidence base for services working with youth at-risk of involvement in the criminal justice system: Developing a standardised program approach.* Health & Justice, 6(1), 8. https://doi.org/10.1186/s40352-018-0066-5

Lee, Ebesutani, C., Kolivoski, K. M., Becker, K. D., Lindsey, M. A., Brandt, N. E., Cammack, N., Strieder, F. H., Chorpita, B. F., & Barth, R. P. (2014). *Program and Practice Elements for Placement Prevention: A Review of Interventions and Their Effectiveness in Promoting Home-Based Care.* American Journal of Orthopsychiatry, 84(3), 244–256. https://doi.org/10.1037/h0099811

McCarthy, S., & Griffiths, L. J. (2021). *The journey to evidence: adopting evidence-based programs in an Australian child welfare organization.* Human Service Organizations: Management, Leadership & Governance, 45(3), 273–280. https://doi.org/10.1080/23303131.2021.1906372

Mccracken, S., & Marsh, J. (2008). *Practitioner expertise in evidence-based practice decision making. Research on Social Work Practice* - RES SOCIAL WORK PRAC, 18, 301–310. https://doi.org/10.1177/1049731507308143

NSW Department of Communities and Justice. (2022a). *The evidence portal.* https://evidenceportal.dcj.nsw.gov.au/evidence-portal-home/using-evidence/what-is-evidence.html

NSW Department of Communities and Justice. (2022b). *Using a core components approach.* https://evidenceportal.dcj.nsw.gov.au/evidence-portal-home/using-evidence/what-is-evidence.html

NSW Department of Communities and Justice, (2022c). *Preventing child maltreatment: Core components.* https://evidenceportal.dcj.nsw.gov.au/evidence-portal-home/our-evidence-reviews/preventing-child-maltreatment/core-components.html

NSW Department of Communities and Justice. (2022d). *Evidence portal technical specifications.* https://evidenceportal.dcj.nsw.gov.au/evidence-portal-home/about-the-portal/evidence-portal-technical-specifications.html

Nutley, S. M., Powell, A. E., & Davies, H. T. O. (2013). *What counts as good evidence [Report].* Alliance for Useful Evidence. https://research-repository.st-andrews.ac.uk/handle/10023/3518

Oxford Languages. (n.d.). *Evidence.* In Oxford Languages Dictionary. https://languages.oup.com/google-dictionary-en

Petticrew, M., & Roberts, H. (2003). *Evidence, hierarchies, and typologies: Horses for courses.* Journal of Epidemiology & Community Health, 57(7), 527–529. https://doi.org/10.1136/jech.57.7.527

Practicewise. (2023). *What works in children's mental health?* https://www.practicewise.com/

Proctor, E. K., & Rosen, A. (2008). *From knowledge production to implementation: Research challenges and imperatives.* Research on Social Work Practice, 18(4), 285–291. https://doi.org/10.1177/1049731507302263

Ryan, G. (2018). *Introduction to positivism, interpretivism and critical theory.* Nurse Researcher, 25(4), Article 4.
SA Department of Human Services. (2022a). Early Intervention Research Directorate. https://dhs.sa.gov.au/services/cfss/eird/common-elements

SA Department of Human Services. (2022b). *Explaining the common elements.* https://dhs.sa.gov.au/__data/assets/pdf_file/0006/129354/Explaining-the-Common-Elements-Background-and-Introduction.pdf

Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). *Revisiting the quantitative-qualitative debate: Implications for mixed-methods research.* Quality & Quantity, 36(1), 43–53. https://doi.org/10.1023/A:1014301607592

Soydan, H., Mullen, E. J., Alexandra, L., Rehnman, J., & Li, Y.-P. (2010). *Evidence-based clearinghouses in social work.* Research on Social Work Practice, 20(6), 690–700. https://doi.org/10.1177/1049731510367436

Van der Put, Assink, M., Gubbels, J., & Boekhout van Solinge, N. . (2018). *Identifying effective components of child maltreatment interventions: A meta-analysis.* Clinical Child and Family Psychology Review, 21(2), 171–202. https://doi.org/10.1007/s10567-017-0250-5

Van Dyke, M., Kiser, L., & Blase, K. (2019). Heptagon Tool. Chapel Hill, *NC: Active Implementation Research Network.* www.activeimplementation.org/resources

Vindrola-Padros, C., Pape, T., Utley, M., & Fulop, N. J. (2017). *The role of embedded research in quality improvement: A narrative review.* BMJ Quality & Safety, 26(1), 70–80. https://doi.org/10.1136/bmjqs-2015-004877

Waters, R. A., & Buchanan, A. (2017). *An exploration of person-centred concepts in human services: A thematic analysis of the literature.* Health Policy, 121(10), 1031–1039. https://doi.org/10.1016/j.healthpol.2017.09.003