

Spotlight

**Informing teaching: navigating and translating education
best practice**

March 2021



Contents

Introduction

Questions to consider

Quality – How supported is this research?

Reliability, validity and design study – How robust is the research?

Sampling – Is the sample appropriate?

Significance – Are the findings meaningful?

Implementation – How can this research be applied in practice?

Conclusion

References

This Spotlight addresses five questions to consider when navigating best practice in education:

- ⚡ **Quality** – How supported is the research?
- ⚡ **Reliability, validity and design study** – How robust is the research?
- ⚡ **Sampling** – Is the sample appropriate?
- ⚡ **Significance** – Are the findings meaningful?
- ⚡ **Implementation** – How can this research be applied in practice?

Introduction

At a time when people have greater access to information than ever before, the ability to distinguish between what is robust research, and what is not, is a vital skill.

Research literacy is especially important in professions where day-to-day decisions have the potential to significantly impact the lives of others. Doctors must ensure the treatments they recommend are the most effective and efficient options available to meet the needs of their patients. Similarly, teachers and school leaders seek to implement the best strategies to maximise student learning. Since action should be guided by evidence, the process of reflecting on how to engage with research and think critically about information can be beneficial for everyone. This Spotlight is focused on a few topics that can guide the interrogation of research papers, scholarly publications, journal articles and conference presentations to ensure decisions are supported by strong evidence.

Teachers and school leaders are responsible for supporting the learning outcomes of students with broadly varying needs and abilities. Every day, education professionals must make complex decisions about classroom strategy, curriculum, assessment approaches and student welfare. Given their importance, strategies implemented to maximise student learning should be evidence-informed.

Expanding the skill to recognise quality research is essential to help teachers and school leaders become better evaluators and consumers of evidence. It can sometimes be difficult to distinguish practices that are robustly supported by research from those that are based on more limited findings. Knowing what questions to ask when engaging with research can help build confidence in using evidence in practice.

Evidence is a contested notion and debates continue about how to best determine its quality. Besides research evidence, educators may encounter forms of evidence in their daily life through classroom observations, talking with their students and reviewing their students' work. These forms of evidence, are legitimate, and can be considered alongside other evidence and triangulated to inform future teaching and learning decisions.

As a starting point, discussing the five questions detailed here will provide teachers and school leaders with an approach for interrogating education research. By asking these questions, education professionals will be better equipped to understand the strengths and limitations of the research they encounter. They will also help educators determine whether the content and context of the research is relevant to their own circumstances, which should help them decide whether to adapt it to their own setting.

Questions to consider

Quality – How supported is the research?

- Have the research findings been replicated in other studies?

- Do the research findings indicate positive effects on learner outcomes and have these effects been demonstrated across different contexts?

Reliability, validity and design study – How robust is the research?

- Is the method/measurement instrument reliable and valid?
- Has the research method controlled for critical factors that may influence the results?
- Does the study design appropriately isolate the phenomenon of interest?

Sampling – Is the sample appropriate?

- Is the sample broadly representative of the population of interest?
- Are the findings generalisable across contexts, like your own school?

Significance – Are the findings meaningful?

- Do the findings indicate there is a true relationship between the groups or factors of interest?
- What is the size of the effect?
- Do the findings have real world significance?

Implementation – How can this research be applied in practice?

- Is the evidence-based practice appropriate for your context and the needs of your learners?
- Which elements of the research should be implemented with high fidelity and which can be adapted to suit your context?

Quality – How supported is this research?

The concept of evidence-based or evidence-informed practice is a scientific method of quantifying what works, for whom, under what circumstances. It is an approach that is adopted in a range of professions and industries. For example, in healthcare doctors, nurses and other health professionals weigh evidence from clinical trials, or other scientific validation of treatment

options, against their own professional experience (Masters, 2018). This is applicable in school contexts: highly effective teachers rely on evidence of students' learning as well as evidence-informed teaching strategies to shape their teaching practices. They also use and generate evidence to understand learners' progress and to ascertain the effectiveness of their teaching strategies.

For teachers, identifying effective classroom strategies to ensure positive outcomes for their learners is critically important. It may be difficult to differentiate practices grounded in solid research from those that are not. Understanding this difference is an important part of engaging with research.

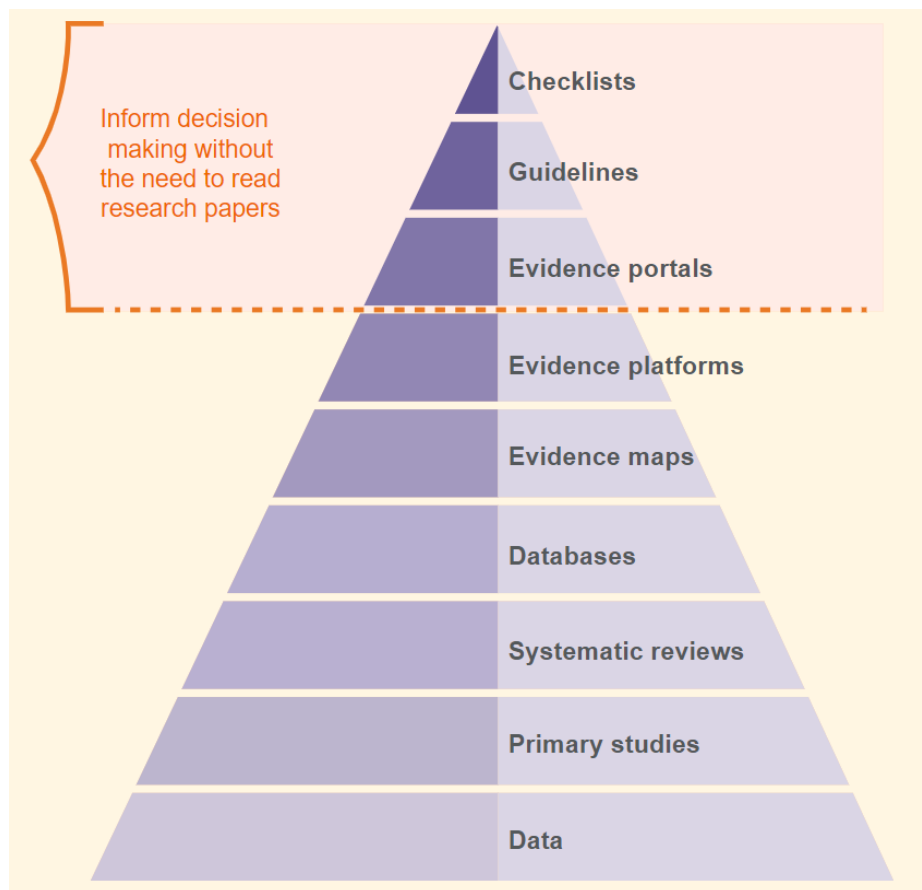
Research-based vs evidence-based and evidence-informed

There is a distinction between research-based and evidence-based practices. For example, a single study investigating the use of a specific reading program in a small classroom may report a positive result. While the reading program may be grounded in research, unless this finding is replicated in different settings it cannot yet be considered 'evidence-based practice.' That is, a single study is insufficient to provide guidance to many teachers on what is likely to work in different contexts or classrooms.

Evidence-based practice refers to findings that have been robustly evaluated and replicated by other researchers, and published in peer-reviewed journals, and where there is a broad consensus within the research community that 'a critical mass of studies that point towards a particular conclusion' (Stanovich & Stanovich, 2003). The effectiveness of the practice or research finding is generally agreed by experts due to thorough, rigorous and repeated demonstration of results.

The concept of *evidence-informed* practice is also important. In education, this reflects the fact that in practice, educators apply their own professional judgement alongside evidence (Sharples, 2013). That is, research is not the sole source of information utilised by educators in their day-to-day decision making and practitioner expertise is also critical (Nelson & Campbell, 2017). Research evidence will never be able to replace educators' professional experience and their unique understanding of their students and the school environment in which they work, but it can supplement this important knowledge (Education Endowment Foundation, 2019).

The evidence pyramid (White, 2020)



Long text description

Data forms the basis of all good research and evidence. As demonstrated in the evidence pyramid above, it can be operationalised into useful guides, checklists and other resources for practitioners to use in their day-to-day practice (top of the pyramid). These resources allow practitioners to utilise strong evidence and incorporate it into their practice without needing to engage with underlying detailed data sets.

Evidence can take many forms. Primary studies collect and report on data generated through an empirical research study, while systematic reviews collate primary studies to present multiple pieces of evidence on a specific topic. Data can be stored in databases, which are useful for comparing sets of data, for example, from different schools or different jurisdictions. Evidence maps illustrate the quantity, distribution and characteristics of published studies and identify gaps in research, highlighting future research needs. Evidence platforms usually exist for specific sectors and guide users to recognise evidence sources. At the top of the pyramid, evidence portals, guidelines and check-lists are often based on data and research findings but do not necessarily reference the research directly.

Assessing the evidence base

There are several ways to find out whether a practice is evidence-based. Literature reviews look at a range of findings related to a specific intervention or practice, using specific criteria to guide evidence evaluation. Such detailed reviews contribute to the evidence base by aggregating,

summarising and critically analysing the effectiveness or impact of the practices in question.

For example, in 2013 the Australian Council for Educational Research (ACER) published the results of a literature review they conducted into the efficacy and effectiveness of various numeracy and literacy interventions in early schooling. The findings demonstrate the differences between evidence-based general principles and individual research-based interventions and teaching practices. While little corroborating evidence was found for many specific practices, ACER identified several general principles essential to the design and implementation of literacy and numeracy interventions. This review offers an important resource that situates specific interventions within the broader context of general principles that should underpin all successful literacy and numeracy teaching practices.

Resources

Literature Review (ACER, 2013)

Read more about ACER's review of literacy and numeracy interventions in the early years of schooling findings.

🔗 https://research.acer.edu.au/policy_analysis_misc/20/

Resources to help educators locate quality evidence

In addition to specific reviews conducted by organisations such as ACER, or research teams at universities, there are some valuable online resources that provide accessible summaries of evidence in education research. These resources assist educators to make evidence-informed decisions to suit their needs through useful snapshots that weigh the impact and cost of various interventions and practices.

Australia

Evidence for Learning

🔗 [Evidence for Learning - Toolkits](#)

UK

Education Endowment Foundation, Teaching and Learning Toolkit

🔗 [Education Endowment Foundation, Teaching and Learning Toolkit](#)

USA

Institute of Education Sciences, What Works Clearinghouse

🔗 ies.ed.gov/ncee/wwc/FWW

Reliability, validity and study design – How robust is the research?

Strong evidence is based on research and evaluation methods that are both reliable and valid. The reliability and validity of methods, tools and measurement instruments can be estimated using statistical analysis. Though related, reliability and validity have distinct meanings in research.

A method or measurement tool is **reliable** if it produces stable and consistent results. For example, a reading test may be considered reliable if it provides the same result for the same student each time it is administered. Similarly, a diagnostic learning tool administered to a single student by multiple teachers may be reliable if it yields the same diagnostic result each time for the same student, irrespective of which teacher administers the tool.

A measurement instrument or method can be highly reliable, but not valid. **Validity** is a judgement of the extent to which evidence and theories support the appropriateness of actions and assumptions based on test scores or other assessments (Messick, 1989). At its core, validity relates to whether the available evidence is relevant to the intended interpretation and appropriate in the given context (Bandalos, 2018). For example, a highly reliable reading test may lack validity if it consistently yields a student reading score several percentiles lower than an appropriately calibrated alternative measurement tool. This type of validity is an important requirement of all scientific research to ensure the method used is an accurate and appropriate way to answer research questions. This is a particular challenge for education research where the concept being measured may be complex and multifaceted like well-being or literacy skills.

In addition to the validity of the measurement instrument, it is also important to consider the validity of the research holistically and how results are interpreted. This is particularly relevant for research aimed at exploring causal relationships between different factors that may be influencing outcomes. If the study detects significant and meaningful impacts, these effects should be attributable to the intervention itself, and not to other factors (Harn, Parisi, & Stoolmiller, 2013). It is also good to be aware if there are multiple interpretations of a particular result or data.

Causation can be investigated in a variety of ways. Running an experiment where the independent variable^[1] is manipulated to investigate the effect on a dependant variable^[2], while controlling for other variables through techniques such as randomisation as is the case in randomised controlled trials, is a methodologically sound way of isolating possible cause and effect relationships. For example, the type of learning intervention delivered might be tested against achievement scores for students who are randomly assigned different learning interventions.

Another method of identifying causation involves correlational studies which look at the effect of non-controlled variables (such as the number of books in the home) on learning outcomes. These types of experiments are known as observational studies, where relationships between variables are observed without the level of manipulation or direct control specific to experimental studies.

Research methods

Different kinds of research answer different kinds of questions. Research that generates quantitative data is often used to measure variables and verify existing theories or hypotheses using statistical methods. Quantitative data can be generated through a variety of tests including experiments, controlled observations and surveys and questionnaires. On the other hand, meanings, beliefs and experiences are often better captured through research that generates qualitative data, for example through interviews and focus groups.

It is often useful to 'triangulate' data, which involves using multiple types of research methodologies to collect data on the same topic to ensure the findings are valid. This is also referred to as using a mixed methods approach. Triangulation allows for both measurement and a deeper understanding of a certain phenomenon of interest. For example, a teacher could survey their students on their preferred mode of learning through an itemised survey with rating scales and then conduct interviews or focus groups with their students to further investigate the findings from the survey data. This method would provide a mix of quantitative (how many students prefer a particular method) and qualitative data (why students prefer one method over another) and offer the teacher rich information about the learning preferences of their students.

When causal relationships are investigated, it is important to consider the validity of the study in making claims about the nature of the relationship. This is called internal validity, and it describes the extent to which an observed effect can be attributed to a causal relationship between the variables investigated, and not due to other factors. Internal validity is particularly important in research studies that investigate the effectiveness of education interventions.

Threats to internal validity include selection bias, which occurs when the sample that is used in the research have characteristics that are skewed in a particular way and therefore not representative of the population of interest. This is problematic as differences that exist between groups may interact with the independent variables, therefore making it difficult to tell whether inherent group differences or the intervention was responsible for the outcome. Other threats include diffusion effects, which occurs when effects from the intervention or treatment spread to the control group (for example, because participants in the control and intervention groups communicate with each other and share information) thereby making it difficult to identify or measure outcomes; and regression effects. Regression towards the mean can obscure the impact of a learning intervention by masking 'true' results with results attributable to chance or other factors. This usually occurs wherever results are contingent upon a complex interplay of many factors – chances are that extreme results will usually be followed by more average ones. For example, a student may perform very badly on a mathematics quiz then a week later score closer to the class average. Their first test result may have been caused by many factors that were not affecting their performance during the second quiz, such as stress or distraction. Their actual skill level may not have changed from test one to test two.

An important feature of any research finding is the extent to which it can be applied to other situations – this is called *external validity*. External validity is influenced by a variety of factors. One of the most important factors that influences both internal and external validity will be discussed in the next section: the participant sample and how representative of the target population it is.

Resources

Australian Education Research Organisation (AERO)

AERO is Australia's independent education evidence body, working towards excellent and equitable outcomes for all children and young people by advancing evidence-based education in Australia. AERO has released a suite of resources aimed at making high-quality evidence accessible and enhancing the use of evidence in Australian education. The following AERO resources can be used in conjunction with this Spotlight to help educators interrogate research and apply evidence in their practice.

Standards of Evidence

The AERO Standards of Evidence establish AERO's view on what constitutes rigorous and relevant evidence. Educators can use the Standards of Evidence to determine the strength of existing evidence on a particular approach in their context.

🔗 <https://www.edresearch.edu.au/standards-evidence>

Evidence Rubric

The AERO Evidence Rubric supports educators to apply the AERO Standards of Evidence in their context. The rubric helps educators evaluate their confidence in the effectiveness of a new or existing policy, program or practice. The rubric also offers implementation guidance appropriate to one's level of confidence.

🔗 <https://www.edresearch.edu.au/evidence-rubric>

Research Reflection Guide

The AERO Research Reflection Guide helps education practitioners and policymakers reflect on what they have learned from reading a piece of research evidence on a particular policy, program or practice. The guide helps you decide whether to implement the approach in your context, and if so, how to do so effectively.

🔗 <https://www.edresearch.edu.au/research-reflection-guide>

To find out more about AERO visit <https://edresearch.edu.au/>.

Sampling – Is the sample appropriate?

It is rarely possible to involve all members of a population in a study. As such, researchers often collect data from a 'sample' of individuals from their target population. The goal of sampling is to draw on a group that is both sufficient in terms of size and is representative of the population of interest so that the chance of detecting a 'true' result is maximised.

Larger samples are not necessarily inherently better. Sample size is highly dependent on the research goal, and on variation in responses, measurements and other data collected. Large samples may be important if the researcher is seeking to understand the experiences of, and differences between, many heterogeneous groups.

For example, the Teaching and Learning International Survey (TALIS) is a long-running, large-scale survey of teachers, school leaders and their educational settings. In the 2018 cycle, TALIS surveyed 260,000 teachers and 13,000 principals in lower secondary schools, in nearly 50 countries (OECD, 2019b). Such a large sample of participants was required due to the complexity of the concepts measured, the range of experiences of teachers and school leaders across many OECD countries, and the size of the overall population (there are perhaps millions of teachers in the OECD).

The TALIS survey uses a highly robust sampling technique to achieve a sample that broadly represents the target population. This involves both a large number of participants and targeted random sampling to ensure the many subgroups within the overall group of OECD teachers are reached (OECD, 2019a).

Ultimately, the 'right' sample size will depend on the aim of the research. Smaller samples can still be valuable, particularly in qualitative research where there may not be a large degree of variation in views or experiences. For example, a school leader may be interested in the preferences of their teachers for accessing professional learning. In this case, the information sought may be more qualitative, and a smaller sample of opinions may suffice. It may also be possible to sample the whole population of interest in a school-based investigation like this.

Without appropriate sampling techniques, a large sample may not be 'better' than a smaller sample, particularly if the larger sample does not reflect the variability of the overall population. For example, a survey of Australian teachers that collected information from a large sample of teachers in metropolitan Victoria and a similar study involving a smaller sample of teachers from all over the country and across different geographical areas may both reveal interesting findings. The Australia-wide study, though smaller, may offer more information about Australian teachers than the larger Victorian-based study that had a more homogeneous, though larger, sample. Both studies have value, but it is important to consider when findings can be generalised across contexts and when they cannot.

If the research involves a sample of the population, what needs to be considered is whether the sample is *representative* of the population. This is important when there is likely variation between subgroups, as is usually the case in education. For example, a study investigating numeracy achievement in an upper secondary school in a metropolitan, wealthy area, may not yield findings that are applicable beyond similar settings. The sample of an experiment or survey is an important factor in determining how generalisable the results are across contexts.

Significance – Are the findings meaningful?

Researchers usually report whether their findings are statistically significant. Though it may be logical to assume this means the results are meaningful in the real world, this may not be true. In a statistical sense, the word significance refers to whether the relationship, or difference, between

two or more groups is unlikely to be due to chance. It is usually reported as a statistical value, called a p-value.

A p-value is a probability which is the result of a statistical test. Small p-values correspond to strong evidence. For example, if $p=0.05$, one would expect the reported result to occur by chance 5% of the time (five out of every 100 times an experiment was run the result would occur by chance). If the p-value is below the predefined limit (such as $p < 0.05$, or $p < 0.01$) results are designated as "statistically significant". It is important to note that statistical significance says nothing about the size of the difference or the usefulness of the finding.

Consider the following hypothetical example. A researcher is interested in comparing gains in numeracy achievement following the administration of a numeracy program specific to 8-9-year-old children – technique A – to a practice currently in use, or the control group – technique B. The researcher designs and implements an appropriate study involving 1,000 learners aged 8-9 of roughly equal achievement levels. Half receive instruction in technique A and the other half receive technique B. The results indicate that, on average, students receiving technique A gained 7 points in numeracy achievement, whereas students receiving technique B gained 5 points, and this difference was statistically significant. By quantifying that the group of children receiving instruction in technique A performed statistically significantly better than students who received technique B, the researcher has confirmed that the result was probably not due to chance. That is, this example demonstrates that statistical significance is a valuable way of checking whether a detected effect was real. In cases where there is a large amount of variation within or between the groups being compared it is especially important to quantify the probability that the obtained results were due to an actual effect rather than some other factor within the particular sample (sampling error).

Statistical significance does have limitations. With large sample sizes, even very small differences between groups or variables can be detected through statistical analysis. How meaningful these differences or relationships are is dependent on the real-world context of the findings. Statistical significance does not provide this information. This is why measuring the size of the difference is important (e.g. the effect size).

In the case of the hypothetical study comparing numeracy programs – techniques A and B – we need to ask some additional questions, beyond whether the detected difference was statistically significant, or unlikely to be explained by chance. Is this two-point difference sufficiently meaningful to enable the researcher to definitively recommend that technique A is better than technique B for improving 8-9-year-old learners' numeracy achievement? Is the two-point gain significant in a real-world sense? Probably not. However, a two-point difference in another context could have a significant impact. For example, when a cut-off score is applied to a proficiency exam, two points may be the difference between pass and fail. So, in this context, a two-point difference can be significant in a real-world sense even though overall, it may be marginal.

The answers to these questions about real world significance also depend on the kind of tool used to measure numeracy achievement before and after the intervention of techniques A and B, that is, the sensitivity of the tool in detecting gains in student achievement, and the magnitude of the effect. In education research it is important to consider not only whether an effect exists, but also the magnitude of that effect as a way of demonstrating real-world meaning. There are many

statistical ways of quantifying the magnitude of an effect (Kelley & Preacher, 2012). No matter which method is used, it is always important to evaluate the real-world significance of research findings.

Magnitude and effect size

Learn more about the magnitude of an effect and how effect size can be used in education research

🔗 <https://www.nzcer.org.nz/research/publications/how-much-difference-does-it-make-notes-understanding-using-and-calculating-eff> (Schagen & Hodgen, 2009)

Implementation – How can this research be applied in practice?

Deciding which interventions may be useful for specific contexts is only half the equation. Once teachers and school leaders have engaged with research and weighed the value of specific interventions for their learners, they next need to incorporate evidence-based practices. Understanding how to implement research-based interventions requires consideration of both the appropriateness of the intervention in the context, as well as the importance of balancing the need to implement the intervention with the requisite fidelity to produce the desired outcomes, and customising the intervention to suit the specific needs of the education setting and its learners (Harn et al., 2013).

Fidelity of implementation is broadly defined as the degree to which an intervention or evidence-based practice is implemented as intended (Harn et al., 2013). Educators are constrained by the “often unpredictable and sometimes chaotic” realities of real-world classrooms with students of varying needs and abilities, specific resource limitations, time constraints, and various other day-to-day circumstances in schools (Harn et al., 2013). These factors (that research studies may not account for) influence an educator’s ability to implement an evidence-based practice with fidelity. When applying evidence-based practices to classrooms, measuring fidelity is a way to monitor the quality of implementation.

The process of understanding how to implement the desired intervention with fidelity can be used as an informative, evaluative process for teachers and school leaders. In fact, the process of evaluating the implementation can be a useful tool for educators by promoting timely and reactive feedback to help address any gaps or problems as they appear (Harn et al., 2013).

When implementing a new evidence-based practice, school personnel should measure fidelity early and often to provide timely and responsive professional development and maximize student outcomes

A key part of this process is determining which aspects of an intervention are the critical components or *active ingredients* that should not be altered during implementation (Harn et al., 2013; Stains & Vickrey, 2017). Critical components can be structural (e.g. materials, timing/frequency of intervention activities) or process (e.g. behaviour/activities of teachers), and will differ depending on the teaching practice in question (Stains & Vickrey, 2017).

While evidence-based practices may have been deemed successful in tightly controlled conditions, implementation in practice will require a degree of flexibility in real-world circumstances. Research suggests that adapting evidence-based practices during implementation may actually increase the success of the intervention and promote its sustainability over time (Harn et al., 2013). A critical step in applying research in schools and education settings is understanding what aspects of the research should be implemented with high fidelity (the critical components), and what components may be acceptably altered to suit real-world contexts. This is where the concept of evidence-*informed* practice is so important – ultimately, an educator’s professional judgement combined with quality evidence are critical ingredients in making teaching practices work best for students. While adaptation is appropriate to ensure a practice is relevant in a particular context, teachers should be cautious in their adaptations, as changes may have an effect on how impactful the practice is as a result.

In recognition of the diverse real-world circumstances of classrooms, researchers should actively seek to aid implementation by identifying the critical components of the practice versus those that are adjustable, such as the location of program delivery or the timing of certain activities (Harn et al., 2013). If the critical components are unclear from research studies, educators may consult with experts or refer to additional research findings to establish those elements that should be implemented with high fidelity versus those that are malleable.

Implementation in schools and education settings

Implementation guide

Read more about Evidence for Learning’s implementation guide for schools.

🔗 evidenceforlearning.org.au/guidance-reports/putting-evidence-to-work-a-schools-guide-to-implementation/

Implementing new learning interventions in schools is a multistage process, rather than a stand-alone ‘event.’ This process should be underpinned by a school environment that is both conducive to embedding the implementation of new practices in day-to-day operations and supported by strong leadership that cultivates a shared approach to implementation. In the Guidance Report: Putting Evidence to Work, Evidence for Learning suggests that these two factors – recognising that implementation is a process, and the role of school leadership and a school environment conducive to implementation – provide the foundation for good implementation (Evidence for Learning, n.d.-a). Once these foundations are in place, and an

appropriate learning intervention has been identified to meet the school's need, implementation should proceed in stages that maximise both the effectiveness of the intervention and the likelihood that it will become embedded in school practice.

The Education Endowment Foundation (2018) has developed a guidance report to help educators utilise research evidence, *Putting evidence to work: a school's guide to implementation*. It makes six recommendations to support educators through the implementation process:



Long text description

Participating educators should receive appropriate up-front professional learning to ensure consistent understanding of the critical components of the implementation plan, thereby maximising their confidence and fluency in the intervention (Harn et al., 2013). Consideration

should also be given to the desired outcomes from the intervention and the implementation plan should consider how these will be measured and evaluated. During the implementation, ongoing coaching and monitoring should be prioritised. By supporting staff at each stage of implementation and monitoring outcomes consistently, the implementation of critical components can be tailored as required to ensure the intervention is adapted appropriately to best suit the needs of learners. When the intervention is working well, good practice should be rewarded and scaled-up to allow the implementation to be sustained (Evidence for Learning, n.d.-a).

Resources

Implementing evidence-informed practice

Some jurisdictions have developed guidance for educators exploring how to implement evidence in the classroom

Australian Capital Territory

Great Teaching By Design

🔗 http://www.det.act.gov.au/_data/assets/pdf_file/0006/854466/Great-Teaching-by-Design.pdf

New South Wales

Centre for Education Statistics and Evaluation's What Works Best

🔗 <https://www.cese.nsw.gov.au/publications-filter/what-works-best-2020-update>

Queensland

Standards of Evidence

🔗 <https://qed.qld.gov.au/publications/management-and-frameworks/evidence-framework/foundations-evidence/standards-evidence>

Victoria

High Impact Teaching Strategies

🔗 <https://www.education.vic.gov.au/Documents/school/teachers/support/Expired/0000highimpactteachstrat-expired.pdf>

The importance of collaboration

Professional collaboration is an important part of everyday teaching practice (Australian Government Department of Education and Training, 2018). This extends to evaluating and using research evidence. When teachers and school leaders come together to discuss research findings, they can pose questions and answers together and develop a shared understanding of

the findings. This shared understanding is crucial for implementation, as it can help to embed the research consistently across the learning environment (Sharples, 2013). In discussing and interrogating evidence together, educators can operationalise research findings to ensure the development of best practice that meets their students' specific needs (Hargreaves & Fullan, 2012).

Each school will have a different approach to collaboration. This approach will depend on the collaborative structures that already exist within the school, such as professional learning communities, or the capacity to develop these. Each school will also need to think about the best approach to adopting evidence-informed practice in their context. Employing a collaborative approach can reduce the burden on the entirety of the school staff to engage extensively with education research. Staff can be introduced to different methodologies, current research debates and issues that are relevant to their unique school community and stay abreast of the latest developments in educational research. Through collaborative activities, teachers can introduce research evidence to their collective practice for the benefit of their students.

Conclusion

Consuming research to stay up to date with the latest advances in the education evidence base is an important and worthwhile practice for all education professionals. While improving research literacy can be challenging, it should not be seen as an insurmountable obstacle to understanding the importance and necessity of evidence-informed practice. Critically engaging with education research, by considering the questions outlined here, will assist teachers and school leaders to distinguish between practices that may be useful for their learners, and those that are unlikely to work in their context. There are numerous resources available for teachers and school leaders to help with this process. Some have been highlighted throughout this spotlight, like *Evidence for Learning's Guidance Report* for implementing research in schools and education settings.

The process of understanding what is likely to work, for whom, and in what circumstances can appear daunting, but it doesn't have to be. By understanding a few simple rules and methods it is possible to evaluate teaching practices and assess their suitability for different contexts. The tips outlined in this *Spotlight* provide guidance on questioning the reliability, validity and generalisability of research, and asking whether the findings are meaningful. Research literacy provides a useful critical framework for teachers and school leaders in implementing the right evidence-informed practices for their school context.

Footnotes

1. Variable that is controlled or changed and assumed to impact the dependent variable.
2. Variable being measured that is 'dependent' on the independent variable.

References

Australian Council for Educational Research. (2013). Literacy and Numeracy Interventions in the Early Years of Schooling: A Literature Review. Report to the Ministerial Advisory Group on Literacy and Numeracy, NSW Department of Education and Communities.

Australian Government Department of Education and Training. (2018). Through growth to achievement: report of the review to achieve educational excellence in Australian schools. Australian Government Department of Education and Training.

Bandalos, D. L. (2018). Measurement theory and applications for the social sciences. Guilford Press.

Bennett, T. (2015). The School Research Lead. Education Development Trust. Retrieved from <https://www.educationdevelopmenttrust.com/EducationDevelopmentTrust/files/93/93c332a4-40df-41ac-8a9b-f803c6573d10.pdf>

Education Endowment Foundation. (n.d.). Teaching and Learning Toolkit. Retrieved December 12, 2019, from <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/>

Education Endowment Foundation. (2018). Putting evidence to work: a school's guide to implementation. Retrieved November 4, 2020 from https://educationendowmentfoundation.org.uk/public/files/Publications/Implementation/EEF_Implementation_Guidance_Report_2019.pdf

Education Endowment Foundation. (2019). The EEF guide to becoming an evidence-informed school governor and trustee. Retrieved June 5, 2020, from https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Guide_for_School_Governors_and_Trustees_2019.pdf

Evidence for Learning. (n.d.-a). Putting Evidence to Work: A School's Guide to Implementation Guidance Report. Retrieved from <https://evidenceforlearning.org.au/assets/Guidance-Reports/Implementation/Guidance-Report-Putting-evidence-to-work-a-schools-guide-to-implementation2.pdf>

Evidence for Learning. (n.d.-b). The Toolkits. Retrieved December 12, 2019, from <https://evidenceforlearning.org.au/the-toolkits/>

Hargreaves, A., & Fullan, M. (2012). Professional capital: transforming teaching in every school. Teachers College Press.

Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing Fidelity with Flexibility and Fit: What Do We Really Know About Fidelity of Implementation in Schools? *Exceptional Children*, 79(2), 181–193. <https://doi.org/10.1177/001440291307900204>

Institute of Education Sciences. (n.d.). What Works Clearinghouse. Retrieved December 12, 2019, from <https://ies.ed.gov/ncee/wwc/FWW>

- Kelley, K., & Preacher, K. J. (2012). On Effect Size. *American Psychological Association*, 17(2), 137–152. <https://doi.org/10.1037/a0028086>
- Masters, G. (2018). The Role of Evidence in Teaching and Learning. Australian Council for Educational Research, 1–5.
- Messick, S. (1989). Validity. In *Educational measurement*, 3rd ed. (pp. 13–103). American Council on Education.
- Nelson, J., & Campbell, C. (2017). Evidence-informed practice in education: meanings and applications. *Educational Research*, 59(2), 127–135. <https://doi.org/10.1080/00131881.2017.1314115>
- OECD. (2019a). TALIS 2018 Results (Volume I): Teachers and School Leaders as Lifelong Learners. <https://doi.org/10.1787/1d0bc92a-en>
- OECD. (2019b). TALIS 2018 Technical Report. Paris: OECD.
- Rai, S. K., Yazdany, J., Fortin, P. R., & Aviña-Zubieta, J. A. (2015). Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. *Arthritis Research and Therapy*, 17(1). <https://doi.org/10.1186/s13075-015-0658-6>
- Ross, W., Vallée-Tourangeau, F., & Van Herwegen, J. (2020). Mental Arithmetic and Interactivity: the Effect of Manipulating External Number Representations on Older Children's Mental Arithmetic Success. *International Journal of Science and Mathematics Education*, 18(5), 985–1000. <https://doi.org/10.1007/s10763-019-09978-z>
- Schagen, I., & Hodgen, E. (2009). How Much Difference Does It Make? Notes on Understanding, Using, and Calculating Effect Sizes for Schools. Retrieved from Research Division, Ministry of Education, New Zealand Council for Educational Research website: <http://www.nzcer.org.nz/system/files/16886.pdf>
- Sharples, J. (2013). Evidence for the frontline. Alliance for Useful Evidence. Retrieved from <https://www.alliance4usefulevidence.org/assets/EVIDENCE-FOR-THE-FRONTLINE-FINAL-5-June-2013.pdf>
- Stains, M., & Vickrey, T. (2017). Fidelity of Implementation: An Overlooked Yet Critical Construct to Establish Effectiveness of Evidence-Based Instructional Practices. *CBE Life Sciences Education*, 16(1), 1–11. <https://doi.org/10.1187/cbe.16-03-0113>
- Stanovich, P. J., & Stanovich, K. E. (2003). *Using Research and Reason in Education*. Portsmouth, New Hampshire: RMC Research Corporation.
- White, H. (2020). Why we need a World Education Organisation. Campbell Collaboration. Retrieved from <https://campbellcollaboration.org/blog/why-we-need-a-world-education-organisation.html>