

From evidence-informed to evidence-based: An evidence building framework for education

Kaydee L. Owen^{1,2}  | Richard C. Watkins^{1,2,3}  |
J. Carl Hughes^{1,2} 

¹Collaborative Institute for Education Research, Evidence, and Impact (CIEREI), Bangor University, Bangor, UK

²School of Educational Sciences, Bangor University, Bangor, UK

³Regional School Effectiveness and Improvement Service for North Wales (GwE), Colwyn Bay, UK

Correspondence

Kaydee L. Owen, Eifionydd and Rhos Normal Site, Bangor University Bangor, Gwynedd, LL57 2HQ, Bangor, UK.
Email: kaydee.owen@bangor.ac.uk

Funding information
None

Abstract

Developing an evidence-base for educational provision can be an arduous journey. To facilitate consistent and accurate implementation of promising approaches, research efforts need to answer questions relating to theoretical coherence, efficacy, effectiveness, cost, ease of use, and acceptability. An evidence-building framework can help us to think critically about the current state of evidence and plan a long-term research programme. In this conceptual paper, we draw attention to an evidence-building framework from medical and psychosocial science that conceptualises the phases and challenges along this journey. We have adapted this framework to help researchers and other education practitioners conceptualise how to review, synthesise, generate and communicate evidence in education. This framework facilitates a *questions-to-methods* approach in our journey to evidence, dissolving the idea that some research and analysis methods are superior to others. We argue that this evidence building framework can help us to distinguish between key terms in education, such as *evidence-informed* and *evidence-based* practice, as well as *teacher enquiry* and *research*—helping to distinguish the roles of researchers and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Review of Education* published by John Wiley & Sons Ltd on behalf of British Educational Research Association

practitioners. This paper has direct relevance to informing how we build evidence in education and identifies a need for collaboration as we work towards an evidence-based teaching profession.

KEYWORDS

collaborative research, educational policy, evidence-based, evidence-informed, translational science

Context and implications

Rationale for this study

With an increasing use of terms such as evidence-informed and evidence-based practice in education, we need guidance to help define these terms more clearly and to structure schemes of research more effectively.

Why the new findings matter

This paper discusses a framework to help educational researchers consider their research questions and methods in a more systematic way. By being critical about the types and quality of research we produce, we can build a more robust evidence-base for an intervention or approach; considering its impact on pupil outcomes, how easy stakeholders find it to implement it in practice, and its cost-effectiveness.

Implications for researchers, policy makers and practitioners

Whilst this paper is written for the intended audience of researchers and teachers engaging in professional enquiry and research, we believe that this information should be shared with educators and policymakers to help them make better informed decisions about intervention buy-in and promotion.

INTRODUCTION

Background and context

Evidence and prosperity in education

To drive prosperity, we need to ensure that every pupil has access to high quality education, and we instil equity into our practices (Berger & Fisher, 2013). This is particularly important in the context of closing the attainment gap between pupils who live in poverty and their peers (Egan, 2016). Heckman et al. (2013) provided longitudinal evidence to suggest that high-quality educational provision can help disadvantaged pupils access further education or training, improve their employment prospects, and significantly reduce the likelihood of them engaging in criminal acts. By investing money and time into interventions that can

improve pupils' skills, knowledge and experience (this is known as *human capital*, see Alika & Aibieyi, 2014), education can have a positive effect on the future workforce and economy (Keeley, 2007).

Many factors affect school provision, including the availability of resources, practitioner preference, prevailing political and social contexts, and the available evidence for the effectiveness of teaching and learning strategies (Slavin, 2020). Decisions about provision should be driven by both teacher experience and research evidence in parallel. Despite the growing acceptance of terms such as *evidence-based* and *evidence-informed* practice in the field of education, there is still no clear guidance around how to design and deliver the most effective education, while ensuring that investment is worthwhile (Education Endowment Foundation, 2021; Hollands et al., 2019). Without a clear framework to inform the accumulation of knowledge between studies investigating the same educational approach/intervention, findings can become disjointed and incoherent, and the trustworthiness of the research can be questioned (Sims & Fletcher-Wood, 2021).

Research methods

There is a contemporary debate within education that some research methods are more superior than others. Wrigley (2018) argued that the current approach to driving evidence-based practice in education focuses too heavily on effect sizes (i.e., the over-reliance on results from randomised controlled trials [RCTs] and meta-analyses). In doing so, this neglects teachers' experiences, pupils' needs and wishes, and an understanding of the social context. Thomas (2020) further identified that the discourse surrounding the use of *evidence-based* (what works) approaches has created a hierarchy in research methods that has unduly influenced education policy.

We argue that a *question-to-methods* approach offers several advantages over an approach that is driven by perceived hierarchy, epistemology or philosophy. First, a question-to-methods approach helps us to conceptualise what evidence already exists and the questions that still need to be answered to improve practice. The process of answering these questions is not bound by a narrow selection of research methods and analysis techniques (dissolving the perceived hierarchy) but draws on the most appropriate methods to answer those questions. This approach recognises that all research approaches can be rigorous if the studies are well designed, and the outcomes are trustworthy. Second, focusing on the use of multiple approaches can help us to gain a holistic picture of how robust a teaching approach is and establish its potential across different contexts. Questions about 'what works' in education need to focus on theoretical coherence, efficacy, effectiveness, fidelity, challenges with implementation, cost and acceptability across different stakeholders. We cannot provide answers across all these domains through a single research method; we may also need to adapt our approach across different school populations and contexts to account for diversity and promote equity. Finally, the decision about whether an approach is evidence-based is nuanced. We need to consider the amount of evidence available, the trustworthiness of this evidence, and how transferable/generalisable the benefits of an approach are likely to be. If we want to encourage the sustained use of a robust approach/intervention, educators and policy makers need to make informed decisions based on all the available evidence.

Encouraging schools to engage with trustworthy evidence

Generating robust and useful research is only part of the challenge. Researchers also need to consider how to translate and disseminate these findings effectively to educators so that

they can integrate best-practice into their classrooms. Willingham and Daniel (2021) explained that teachers may not be aware of the available evidence because they do not have the knowledge, time or desire to read it. This highlights the importance of adopting a common approach to make the ‘key messages’ from research accessible and easy to interpret, including the consistent use of terminology and an explanation of key concepts. Researchers also need to ensure that they are investigating the benefits of an intervention when minimal implementation support is available following training (i.e., conducting effectiveness and maintenance studies). Communicating the outcomes associated with this type of research is important as it is the closest we will get to understanding the benefits of day-to-day use of an intervention.

Initiatives such as those driven by the Education Endowment Foundation and the Institute of Education Sciences have aimed to mobilise process evaluation and make research findings more accessible to education professionals. However, without careful development of summary metrics (such as efficacy, effectiveness and cost effectiveness), the approach lacks the detail necessary for educators to implement individual interventions effectively. Further, meta-synthesis combines outcome data across different phases of schooling and educational contexts, which can reduce the perceived relevance of the results (Edovald & Nevill, 2021; Higgins, 2020). Willingham and Daniel (2021) explained that some teachers believe that research does not directly translate into their classroom—particularly if their school, or the pupils that they work with, differ significantly from the research context. The use of meta-metrics to summarise an evidence-base has limitations as they often fail to report sufficient detail about the *type* of contributing research. Instead, we should be reporting enough context alongside outcome data to allow teachers to draw their own conclusions about the comparative results they could expect in their own classrooms.

A working framework for evidence-based practice

Gorard (2020) indicated that few publications report the best methods for getting evidence into education. In this paper, we propose that researchers should adopt a systematic framework to help identify gaps in existing knowledge and move from evidence-informed to evidence-based interventions. Researchers should draw on the strengths of different research methods to develop our understanding of how interventions or educational approaches function across different contexts and beyond the termination of funded research efforts. This framework can help researchers to think critically about well-designed research to enhance the overall trustworthiness and utility of their findings.

Thorncroft et al. (2011) presented a model to understand how different types of research evidence contribute to a journey towards developing an evidence-base for medical and

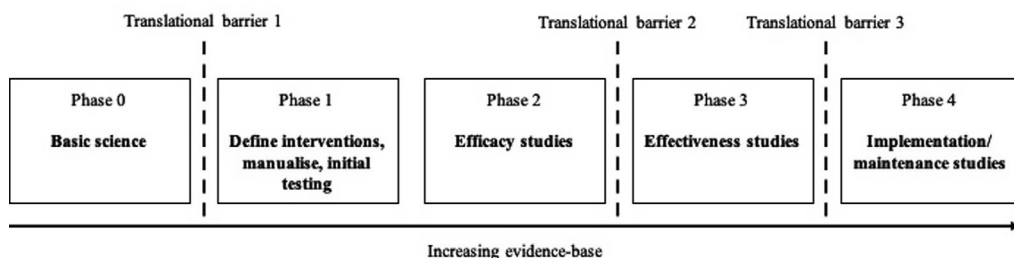


FIGURE 1 An adaptation of Thorncroft et al.'s (2011) model depicting the phases of research that contribute to a growing evidence-base for an intervention

psychosocial intervention (see Figure 1). This journey is characterised by five phases of research and three implementation barriers that make the effective translation of evidence into practice more challenging.

Within *phase 0 (basic science)*, researchers build theories, generate hypotheses, and conduct some small-scale studies to identify initial confounding variables. In some cases, the results do not translate to the wider population (*translational barrier 1*). Within *phase 1 (define, manualise, and test intervention)*, researchers conduct studies that identify the key components of an intervention and begin to define and manualise it. Once an intervention has been clearly defined, researchers can begin some larger-scale exploratory studies. *Efficacy designs (phase 2)* allow researchers to evaluate the outcomes of an intervention or approach under optimal circumstances (i.e., when implementation fidelity is high, and the environment is conducive to success). Larger efficacy trials can also enable researchers to compare the intervention with an appropriate alternative. *Translational barrier 2* indicates an interface between the outcomes of research from efficacy designs and those conducted under routine real-world conditions. *Effectiveness studies (phase 3)* enable researchers to establish whether an intervention can still yield positive results in real-world contexts (where perhaps fidelity is reduced and/or research support for the intervention has been withdrawn). Sometimes the intended stakeholders do not adopt or maintain the intervention after the termination of a research study, despite the evidence-base for that intervention (*translational barrier 3*). The final phase of research, *phase 4 (implementation/maintenance studies)*, identifies factors that interrupt the uptake of evidence-based interventions after the termination of a research study. This includes factors associated with the intention to implement the intervention, early implementation and the persistence of implementation.

The progress from and between phases in Thornicroft et al.'s (2011) model can be circular rather than linear. This is not dissimilar to the typical learning cycles we find in models of close-to-practice research and enquiry (British Educational Research Association, 2018a; Drew et al., 2016). We propose an adapted version of this framework that may help key stakeholders to conceptualise the journey involved in evidence generation for educational provision. This process captures different phases of research from evidence review through to answering key questions about curriculum development, impact, contextual fit, cost, and implementation barriers.

By using this framework alongside a questions-to-methods approach, we can begin to think about the type of research needed to build a compelling argument for approaches that are likely to be effective at scale across an education system for a target population, and how we might mitigate some of the challenges in closing the research-to-practice gap. This framework allows us to work collaboratively and move beyond the question 'what is the quality of this study?' towards 'what is the quality of this study and how does it contribute to the evidence-base?' Systematic progression can help us to identify the best approaches to support our learners, driving prosperity and enhancing equity.

Below we outline how we can adapt Thornicroft et al.'s (2011) framework to plan and synthesise evidence generation within education. Alongside outlining some of the key considerations for each phase, we also provide key questions that practitioners can ask themselves as they engage with research. We have provided an accessible summary of each stage to help non-specialist audiences understand the concepts and ideas. Our aim is to add to the discussion that research generation is a dynamic process, and there are a range of questions that we need to address to enable stakeholders to make well-informed decisions about different teaching approaches (including intervention use). This framework allows us to conceptualise clear differences between evidence-informed and evidence-based practises within education.

THE EVIDENCE-BUILDING FRAMEWORK FOR EDUCATION: CONCEPTUALISING EVIDENCE REVIEW AND GENERATION (ADAPTED FROM THORNICROFT ET AL., 2011)

Phase 0: Basic science

A definition for education research

At the beginning of the research journey, it is important to understand the state of the existing literature. Once a researcher has identified a promising approach/intervention, they should consider which research questions still need to be addressed. They may opt to take a more systematic approach to scope and synthesise the current evidence in relation to a narrower topic. Krnic Martinic et al. (2019) outlined that researchers undertaking a systematic review should report:

- A research question
- The sources they used to undertake their searches (e.g., databases and search engines)
- The inclusion and exclusion criteria for papers
- How they screened each paper
- A critical appraisal of the quality of the included papers.

As such, systematic reviews provide an important starting point for summarising the status of knowledge in a particular field and identifying where evidence is missing (White, 2019). Existing international databases allow researchers and educators to search for published systematic reviews with ease. Examples include the International Database of Education Systematic Reviews and the Campbell Collaboration.

Meta-analyses allow us to synthesise and summarise the quantitative findings from multiple studies. This can be useful in contexts where several studies have assessed the same thing with different learner populations or within different geographical contexts, using similar outcome measures. The results from a meta-analysis such as this may indicate how robust an approach/intervention is in its current form, and/or highlight the need for further research. In their guidance around using meta-analysis for research in education, Pigott and Polanin (2020) suggested that the searches include unpublished data from sources such as dissertations and reports from independent research firms. This helps overcome some of the challenges associated with publication bias, including the tendency for people to only publish research that demonstrates large and statistically significant effects (Polanin et al., 2016).

If a fit-for-purpose intervention does not exist, then we can consider the wider principles (i.e., how children learn or behave) that can inform the logical development of a new package. The process may be underpinned by the core beliefs and values of the researcher—narrowing the literature search to the lens of a learning theory such as constructivism, behavioural science or cognitive science (Ertmer & Newby, 2013). Others may be looking to draw from theories/principles across several fields of learning science. For example, Rosenshine's (2012) ten principles of effective instruction draw from both cognitive and behavioural science. These include aspects of design such as presenting material in small steps, instilling checks for understanding and monitoring independent practice.

Accessible definition for education <i>Phase 0: Basic science</i>	Questions we should ask ourselves	
	Researcher	Teacher
<p>Before developing or adopting an approach/intervention, we should ensure that it is based on sound logic and/or promising theory. In instances where a promising approach/intervention already exists, we can look critically at the existing empirical research to identify the current state of the evidence-base and identify gaps to drive future research.</p> <p>If a suitable approach/intervention does not exist, then researchers and/or curriculum designers can consider how relevant learning and pedagogical theory can help to develop a new one.</p>	<p>Does a promising approach/intervention already exist? If yes, consider how you may develop the evidence-base—what questions still need answering? If not, consider which learning and pedagogical principles you need to incorporate as you develop a new one.</p>	<p>Does the approach/intervention make sense? Is it based on a theory you have previously heard about? Has someone with expertise in the area supported the development of the approach/intervention?</p>

Translational barrier 1

A definition for education research

Sometimes, promising approaches/interventions do not have the desired uptake in real-world classrooms, or they may not have the desired effect when used to support learners. When ideas identified in *phase 0* fail to gain traction, they may need to undergo an iterative process of reflection and refinement to improve aspects of the design and/or contextual fit of the approach/intervention. However, if modifications prove unsuccessful then it might be necessary to return to phase 0 and identify and develop a more promising approach.

An additional barrier that researchers may face is a lack of the financial means or stakeholder buy-in to develop the next phase of research. Research funding is competitive and there is a need to prioritise research that has contextual and political importance at the time of call. Even when researchers are successful in acquiring the funding to conduct a research study or develop an intervention, there is the additional challenge of recruiting schools. If the educational approach is not something that aligns with a school's values or priorities, then recruitment can be challenging, and progress is hindered.

How might we mitigate this barrier?

Logic models allow us to conceptualise how research contributes to the development and uptake of a teaching approach or intervention. They provide a way for research to denote the expected short-, medium- and long-term impact of their work, and can also be used to plan research that spans phases 0–4 of this framework (Coldwell & Maxwell, 2018; Taylor-Powell & Henert, 2008). Logic models also provide space for researchers to acknowledge important assumptions that are imperative to the success of evidence-generation and sustained uptake. This includes inputs, expected outcomes and external/extraneous factors. Connolly et al. (2017) provided a template that researchers in education can use to present their own logic model (Figure 2). If more researchers were to communicate their theory and plans using logic models, we could take a positive unified step towards communicating promising approaches/interventions to key stakeholders (e.g., funders, teachers, policy makers).

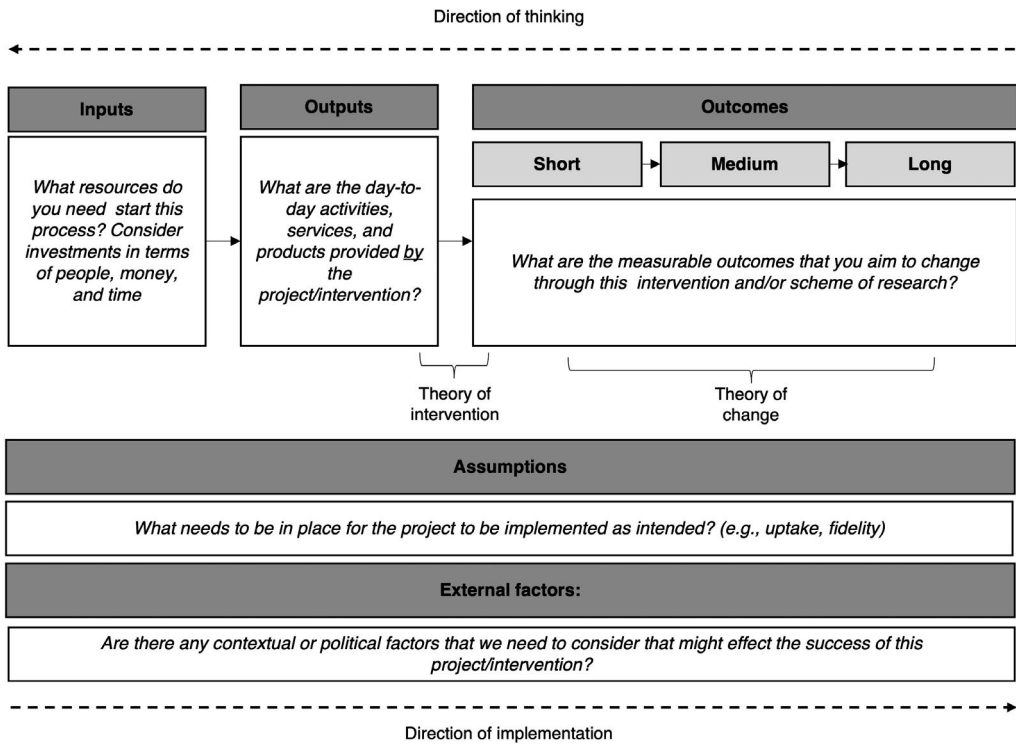


FIGURE 2 A template for a logical model with the key questions we should be asking ourselves as we plan research studies/schemes of work. Adapted from Connolly et al. (2017, pp. 32–38)

Accessible definition for education	Questions we should ask ourselves	
	Researcher	Teacher
Translational barrier 1		
Sometimes, what we think will happen in theory does not materialise in educational settings. This could be because the theory is faulty or the approach does not get it 'off the ground' (i.e., we cannot get funding to support it, or people do not have a need to use it).	Is there a demand for a new approach/intervention? Is the development of this new intervention timely in terms of social and political context(s)? What is the systematic long-term research plan for this intervention?	Is there a need for a new approach/intervention in your school? Will it be practical to use this approach/intervention at scale in your setting? Can you continue to fund this approach/intervention if the cost is not subsidised beyond an initial trial period?

Phase 1: Define, manualise and test interventions

A definition for education research

If the underlying theory is sound and appears to be true in practice, then researchers, teachers and/or designers can begin developing the key components that make up the intervention. This process involves considering the desired overall outcomes, the requirements of in-built tasks, and the learning design necessary to make each of those tasks effective (Eagleton, 2017). Examples of important learning design elements include guided practice, juxtaposing examples, fading scaffolds, mastery assessments and

opportunities to practise building fluency (Binder, 2003; Engelmann & Carnine, 1982; Layng, 2018; Rosenshine, 2012).

To ensure that an intervention or teaching approach is robust, this process often involves iterative phases of field testing (or ‘test of concept’ research). Looking at the literature for an example of this in practice, curriculum designers often ask teachers to test lessons with small groups of pupils to validate that each task is worded in a way that is not ambiguous. They ask teachers to provide detailed feedback about any problems that pupils have faced whilst engaging with the whole programme. This feedback feeds into the revisions before the designers release it for further field testing (Huitt et al., 2009; Koziuff et al., 2000). Only when teachers and pupils report no further problems with the programme, and the questions and instructions are clear and unambiguous, should we begin to systematically test outcomes on a larger scale.

It is important to note that teaching approaches/interventions are often defined and manualised with a specific population in mind. We may have to revisit this phase to make reasonable adjustments for learners with different learning needs, for different settings, or to support different modalities of delivery. It may be helpful to think of an example in the context of the Response to Intervention (RtI) framework, whereby teachers deliver interventions to a whole class, small group or on a one-to-one basis depending on learner needs (Hughes & Dexter, 2011). Typically, the learning needs of children in a school operate on three broad levels: the universal (what do we do for all pupils?), targeted (what do we do for pupils who need more support, perhaps in small, focused groups?) and individualised (what do we do for pupils who might need intensive one-to-one learning support?). If an intervention has been manualised to support individualised delivery, we cannot assume that it will work the same way on a universal scale without adjustments. This also highlights the need to think of these models and stages as cyclical and iterative, rather than simply linear.

Accessible definition for education <i>Phase 1: Define, manualise, and test interventions</i>	Questions we should ask ourselves	
	Researcher	Teacher
During this phase, we begin developing and field testing an evidence-informed approach/intervention, which is underpinned by basic science and effective instructional theory.	Are the instructions, questions, and/or tasks unambiguous? Are any assessments reliable and valid? Do you have early evidence to suggest that the package will have promising results on the specified scale?	Are you able to trial the approach/intervention and provide the designers with feedback to support its development?

Phase 2: Efficacy studies

A definition for education research

Efficacy studies allow researchers to evaluate the outcomes associated with using the approach/intervention. These studies are characterised by a researcher or specialist delivering the approach/intervention on a session-by-session basis to ensure high-levels of implementation fidelity (i.e., in line with the manualised design established during phase 1; Lee et al., 2009). In some cases, a teacher may adopt the approach, with regular guidance from a specialist to support implementation. Well-designed efficacy studies can help us to answer the following questions:

1. Does the approach/intervention have the desired positive effects when it is delivered as intended and optimal support is on offer?

2. How does an intervention compare to a competing approach, when delivered with optimal support?

Answering these questions is an important step in building the evidence-base for an approach/intervention in applied contexts. Different research methods can help us to understand questions around the efficacy of an approach/intervention including randomised controlled trials (RCTs), group comparison studies (e.g., quasi-experimental), and small *N* designs. The questions and contextual restraints should drive the choice of research method.

Group designs allow researchers to assess the impact by comparing the outcome measures of two groups. For example, if a specialist teaches pupils in one group using an intervention, a comparison group would continue to attend their usual classroom lessons without additional support (this is known as a teaching as usual control group). Alternatively, pupils in the comparison group may use a competing intervention (e.g., one that is more popular or is cheaper). The pre- and post-intervention data can help researchers to draw conclusions about which approach leads to the most favourable outcomes.

An RCT is an example of a group design but is arguably a more robust extension. The defining characteristic of an RCT is that a researcher randomly allocates pupils to a research group (Connolly et al., 2017). If enough pupils participate in the trial, the randomisation element aims to make the demographic and attainment profiles of the groups as similar as possible at the beginning of the study (Reich & Milstone, 2014). Assuming other variables have remained constant, we can then attribute any differences between baseline and follow-up outcomes to the intervention itself (Kraft, 2020).

As we move from an evidence-informed approach to one with an evidence-base, researchers may want to consider the quality of their research design and the integrity of their reporting more critically. Gorard (2014) provided some useful criteria to help us judge the trustworthiness of the outcomes of group design research; assuming that the research method is appropriate for the questions and sample. To score the highest rating for trustworthiness (4*), researchers should employ a fair design for comparison, include and retain a large sample of participants, use a standardised and independent measure of outcome, measure and report high levels of implementation fidelity for the intervention across participants, and overcome any potential threats to validity.

However, it is important to stress that not all efficacy trials will lend themselves to group designs. Small *N* research can help us to assess efficacy in instances where group design may not be appropriate or feasible. This design can be useful in education when exploring the impact of individualised interventions or when the population of pupils who will benefit is small (Barger-Anderson et al., 2004; Barnett et al., 2012; Kubina et al., 2002). Graham et al. (2012) explained that small *N* designs typically involve observing pupils' behaviour or attainment scores before, during and after engagement with an intervention, where learners act as their own control. We can then look at the data within and across the different time points to draw conclusions about the impact of the intervention on the outcome of interest. Even outside of a research context, the formative data collected through small *N* design can help to make real-time decisions about the intervention/teaching approach we should use (Lanovaz et al., 2019). Small *N* studies rely on replication rather than large sample statistics to establish reliability and enhance generalisability across populations and settings.

Tate et al. (2008) developed a validated scale to facilitate the appraisal of literature reporting the outcomes of Small *N* studies. The single-case experimental design (SCED) scale consists of 10 items that assess methodological quality and statistical analysis. An additional item assesses whether researchers have reported relevant details regarding participants' clinical history; this allows readers to determine how applicable the findings might be to learners with whom they work.

There are, of course, examples where efficacious implementation occurs beyond the context of efficacy research trials. External providers such as educational psychologists,

speech and language therapists, counsellors and certified behaviour analysts often work in schools to provide individualised intervention, or small group programmes, for those with identified need. The individuals who deliver these specialised services are trained to deliver evidence-informed approaches, so can use them to high levels of fidelity (Barrio Minton et al., 2014; Boyle & Kelly, 2017; Law et al., 2017; Slocum et al., 2014). Efficacy research trials can help us to better inform this type of professional work and highlight which interventions might benefit pupils with varying and often complex needs.

Accessible definition for education <i>Phase 2: Efficacy studies</i>	Questions we should ask ourselves	
	Researcher	Teacher
Once we have developed a new approach/intervention, we need to investigate whether it has the desired effect. Efficacy studies involve a specialist delivering the provision or working closely with a teacher to ensure high-quality implementation.	What elements of implementation are integral to the programme's success? How are you going to measure fidelity throughout the study? Have you adopted an appropriate research design to assess this in practice with the target population?	If an intervention has promising results from efficacy research, have you got the means to employ someone who is highly skilled at delivering the approach/intervention to implement it 'by the book' (e.g., an educational psychologist)? If not, is it possible to up-skill a teacher/teaching assistant within the school to deliver it this way?
Some efficacy studies involve comparing the outcomes of one provision against a competing approach.		

Translational barrier 2

A definition for education research

Translational barrier 2 represents the interface between efficacy and effectiveness research, where the latter is characterised by studies designed to evaluate the impact of interventions under everyday school conditions, and without support from researchers or experts. Due to lower levels of ongoing support and the difficulties of integrating a new, or adapted, intervention into school life, we often see lower reported effect sizes in effectiveness research (phase 3) compared to efficacy research (phase 2). Researchers do not always know which components of an intervention are responsible for the change that we intend to support, as it is often not one factor that supports change, but rather a combination of several (Lee et al., 2009). Therefore, it is important that key stakeholders implement approaches/interventions with fidelity to the manualised design (phase 1) if they are striving to match the impact reported in efficacy trials (Arthur & Blitz, 2000).

How might we mitigate this barrier?

There is a growing body of research to demonstrate the advantages of providing teachers with follow-on mentoring/coaching opportunities to support effective implementation. Initial training provides teachers with the pedagogical and/or programme knowledge that they need to implement an intervention in their classroom. However, there is a need to provide continual professional development opportunities for teachers to support them in the successful integration of this knowledge in practice (Education Endowment Foundation, 2019; Kraft et al., 2018). Following initial training, one-to-one coaching can support teachers to improve the accuracy of their implementation through multiple observations, feedback cycles,

and modelled sessions in their classrooms (Fletcher et al., 2012; Kretlow & Bartholomew, 2010; Owen et al., 2021; Sailors & Shanklin, 2010). Having a member of school staff who is a specialist in intervention delivery could be a more cost-effective and close-to-practice solution to support implementation in classrooms.

Researchers should also consider a long-term plan to upskill staff to implement approaches/interventions with a promising efficacy evidence-base. This can be achieved by integrating training for these approaches into professional development schemes for accredited professions such as educational psychologists, speech and language therapists, child counsellors, or certified behaviour analysts.

Accessible definition for education <i>Translational barrier 2</i>	Questions we should ask ourselves	
	Researcher	Teacher
<p>Sometimes schools need to buy-in specialised support to run and sustain an intervention, rather than adopting a teacher-led approach. This becomes a barrier if a school wishes to scale up an intervention and/or they do not have the financial means to support it.</p> <p>If a teacher tries to deliver an approach/intervention themselves, they may not see the same outcomes as those reported in efficacy studies. This may be because the approach/intervention is not implemented to a comparatively high standard. Unexpected outcomes may lead them to reject the approach/intervention.</p>	<p>Can a team of specialists support the implementation of this intervention at scale after the termination of research efforts?</p>	<p>Does your school have the financial and technical means to support the intervention to its full potential without ongoing support from the research team? (e.g., consider the need to out-source ongoing training and professional development for internal staff, or buy-in specialists to implement the provision on a 1:1 or small-group scale).</p>

Phase 3: Effectiveness studies

A definition for education research

Effectiveness trials focus on measuring impact under everyday classroom conditions. A researcher may run and/or oversee the training and assessment phases of these trials, but they are not *in situ* for each of the intervention sessions. This may mean that fidelity of implementation is often not as high compared to efficacy trials or at the levels intended for the approach/intervention (as described in *phase 1*). The results from this phase help researchers approximate how an approach/intervention works in the 'real-world' if schools continue to invest in it after the study ends.

This phase also captures important aspects of teacher professional enquiry. The term *professional enquiry* is often used synonymously with terms such as *action research*, *reflective practice* and *teacher research* (Elliot, 1991; Kelchtermans, 2021; MacLean & Mohr, 1999; Schön, 1983; Wyse et al., 2021). Researchers at the Collaborative Institute for Education Research, Evidence and Impact (CIEREI) have developed a schema to distinguish between the terms *enquiry* and *research* and that these fall on a *teacher enquiry-to-research continuum*. This clarification can help teachers conceptualise their role in the creation and mobilisation of evidence and how and when they might collaborate with researchers to create skilled teams that are able to apply a range of robust and relevant research tools to help evaluate and improve educational practice (Kelchtermans, 2021). In this schema, *research* involves university collaboration, requiring external ethical approval, and adopting more robust research methods. *Enquiry* projects are typically teacher-driven, contextual, learner focused and close-to-practice. In effect,

enquiry involves teachers undertaking reflective practice in their own classrooms to investigate approaches/interventions that may address key priorities within *their* school.

Willingham and Daniel (2021) explained that one barrier that prevents teachers engaging with research is the perceived lack of relevance of the findings. Enquiry provides one avenue to evaluate an approach/intervention in their own classrooms and improve the sense of relevance to teachers. This also allows teachers to evaluate the impact of an approach/intervention if they need to make reasonable adjustments to it (e.g., does an intervention work if pupils only use it twice a week instead of the recommended three?).

Accessible definition for education <i>Phase 3: Effectiveness studies</i>	Questions we should ask ourselves	
	Researcher	Teacher
Effectiveness research allows us to evaluate the outcomes of an approach/intervention when it is used in everyday educational settings without support from a specialist. This can take the form of university-led research, teacher-led enquiry or a collaboration of researchers and practitioners.	<p>If an approach/intervention has promising results during the efficacy phase, will it produce similar results when you reduce the available implementation support?</p> <p>What are the outcome variables you want to capture (including social validity and fidelity)?</p>	<p>Can you draw on your own professional enquiry skills to evaluate an approach/intervention?</p> <p>Can a teacher, or teaching assistant, in your school deliver the approach/intervention?</p> <p>Could an external research specialist help you evaluate the impact of this approach/intervention?</p>

Translational barrier 3

A definition for education research

Sometimes stakeholders do not adopt an approach/intervention when research efforts end, even if it has a promising evidence-base. Although it is not an exhaustive list, there are several reasons schools may opt to stop using an approach/intervention after a research study ends, including:

- The financial implication when research funding does not subsidise the ongoing costs of the intervention (O'Reilly et al., 2018)
- A view that the intervention package is not flexible enough to be effective in different classrooms or with various student populations (Egan et al., 2019)
- The emergence of competing or more novel interventions (Kraft, 2020; Stoa & Chu, 2020)
- Limited support from the school's senior leadership team to provide provision to integrate the intervention into the day-to-day life of the school (Flannery et al., 2009; Meyer & Behar-Horenstein, 2015), and/or
- Changes in guidance about the focus of interventions or the interventions schools should use, such as the Department for Education's (2022) list of validated phonics programmes.

In a recent study, Pegram et al. (2022) identified 138 different interventions in use across a cluster of 10 schools—67% of these interventions had no published research evidence to support them and only 11% had promising evidence of positive causal impact on pupil outcomes. This poses important questions around why some schools adopt and use interventions with a limited evidence-base, and whether it is even practical to implement multiple interventions to a high level of fidelity. Despite being informed about the quality of evidence-base for these interventions, the schools in Pegram et al.'s study did not make meaningful changes to the

provision they offered (i.e., nearly all decided to persevere with interventions with little to no evidence-base). This raises important questions for future research regarding the process by which leaders and practitioners choose interventions and teaching approaches in schools.

This implementation barrier also captures the financial implications that interventions carry for schools. This includes the economic cost associated with aspects such as training, on-going support, resources (e.g., manipulatives, textbooks, licences), and staff time (e.g., if a teaching assistant is required to work with pupils in small groups or one-to-one). It is important to be transparent about the cost of an approach/intervention before educators adopt it, to enable them to make better informed decisions (Shand & Bowden, 2021). Economic evaluations help mitigate translational barrier 3 by providing a breakdown of direct and indirect costs associated with delivering educational provision (Detrich, 2020). Kraft (2020) outlined that costs matter when evaluating the policy relevance of effect sizes from efficacy and effectiveness research. This is because policy makers consider the return per monetary unit that they invest and the overall cost of an intervention (i.e., how well does something work relative to the cost per pupil *and* the total financial investment required to have the infrastructure in place to scale up?).

How might we mitigate this barrier?

Gorard (2020) identified that disseminating high-quality evidence in a more accessible format via a respected and trusted conduit is likely to be a key feature for mobilisation. We therefore need to consider a more coordinated and unified approach that encompasses research generation, dissemination and effective support to integrate effective practice in schools. Key stakeholders in this process involve researchers, policy makers, trusted conduits, senior leaders and classroom teachers.

We can look to North Wales for an example of how a trusted conduit has been established in the education system. Since its formation in 2013, one of the aims of the Regional School Effectiveness and Improvement Service for North Wales (GwE) has been to mobilise the use of evidence-based provision in schools. This led to a formal partnership with Bangor University and the creation of the Collaborative Institute for Education Research, Evidence and Impact (CIEREI), and has further enhanced opportunities for collaborative research with schools and informed regional policy making (Tyler et al., 2019). [Corrections made on 5 April 2022, after first online publication: 'CollaEREI' in the previous sentence has been corrected to 'Collaborative Institute for Education Research, Evidence and Impact (CIEREI)' in this version.] In this example, GwE is well positioned to fulfil the role of the trusted conduit to help mitigate translational barrier 3. Other countries should consider whether they have an existing regional network that can fulfil a similar role.

Accessible definition for research <i>Translational barrier 3</i>	Questions we should ask ourselves	
	Researcher	Teacher
After trialling an approach/intervention, a school may decide to stop using or investing in it, even if it has a promising evidence-base.	<p>How cost-effective is this intervention, and is a better alternative available?</p> <p>How can you disseminate an accessible summary of the evidence underpinning this approach/intervention to convince stakeholders of its merit? This summary should be transparent about the trustworthiness of the available evidence and include information about cost and effective implementation.</p>	<p>How cost-effective is this approach/intervention? Is a better alternative available?</p> <p>Has someone summarised the evidence supporting this approach/intervention?</p> <p>Consider <i>how</i> you would like to deliver the approach/intervention and <i>which pupils</i> you aim to support. Does the evidence show that it has worked in a similar setting to the one you work in?</p>

Phase 4: Implementation/maintenance studies

A definition for education research

Implementation/maintenance studies aim to identify and report factors that influence the uptake of an approach/intervention outside of a research context. By gaining insight into how an approach/intervention works in the classroom, we can better understand: (1) how to support schools with any barriers to implementation that they face; and (2) if there are any elements that still need to be refined through the manualisation phase (*phase 1*). We can also receive valuable feedback about any positive elements of the design and capture examples of good practice.

Qualitative research methods can be useful research tools in education to help us gather information about stakeholders' thoughts on, and experiences of, an intervention (Braun & Clarke, 2013). Methods include open-ended surveys, focus groups and one-to-one interviews. When thinking about the trustworthiness of qualitative research, Shenton (2004) summarised four criteria that we should consider in relation to the design and reporting of research—credibility, transferability, dependability and confirmability (the quantitative counterparts to these are internal validity, generalisability, reliability and objectivity, respectively). This provides a framework for interpretation of results that may be biased. One of the pertinent points to note from Shenton's paper is that strong qualitative research incorporates triangulation. In education research, a researcher may use different research methods within one study to compensate for the limitations of singular methods and draw on the benefits of others. For example, a study may incorporate questionnaires and focus groups to gain breadth and depth of people's views, and this may be followed up with observations to validate what these findings look like in practice. Triangulation may also allow us to corroborate the experiences and needs of different stakeholders such as teachers, pupils and their parents/guardians. The final type of triangulation involves investigating whether experiences vary across settings (e.g., the recruitment of schools representing learners from varying socio-economic backgrounds). Capturing different perspectives enables researchers to better understand the advantages of, and/or barriers to, implementation issues, and provides useful directions for future work.

Some research methods might be more appropriate than others depending on the sample we are working with. For example, Gibson (2012) argued that researchers should design studies that account for children's developing social competencies. This includes considering their cognitive, linguistic, social, and psychological tendencies/contexts. These aspects of development will determine a child's ability to understand the task a researcher is asking them to engage in. In a qualitative context, a child's developmental competencies will affect their ability to express their thoughts and experiences. Kortessluoma et al. (2003) explained that using interviews with children can be a favourable method of collecting opinion/experience data over surveys because it allows them to have space and time to discuss their views and experiences, without limiting their responses to a narrow range of categories. Moreover, during interviews, pupils' responses are also not bound by their written comprehension abilities.

This phase of the framework can include close-to-practice research, enquiry or impact case studies; where researchers, teachers and other stakeholders reflect critically about the gaps in their knowledge and practice (Tyler et al., 2019; Wyse et al., 2018, 2021). Together, they can co-construct resources or make reasonable adjustments to an existing approach/intervention to meet the needs of their institutional priorities and the pupils they support. They can then work to evaluate a new or adapted provision. This type of research can help schools to engage in the creation, and evaluation of, an approach/intervention that is

relevant to their context. It will give them the confidence to invest in something that is sustainable because it works for them.

Accessible definition for education <i>Phase 4: Implementation/maintenance studies</i>	Questions we should ask ourselves	
	Researcher	Teacher
To help understand some of the benefits and challenges key stakeholders associate with using the approach/intervention, it is important that we capture their views and experiences.	Does the evaluation element (e.g., interviews, focus groups, survey) consider pupils' developing competencies?	Have you shared your experiences of the approach/intervention with those who are able to develop and promote it further?
If an approach/intervention needs considerable adjustments to help it function in a unique context, it may be possible to commission a collaborative piece of work (i.e., whereby stakeholders can work with researchers to co-develop and evaluate an approach).	Can you corroborate findings using triangulation of different research methods or stakeholder groups? What are teachers' and pupils' experiences of the approach/intervention?	If the approach/intervention needs considerable adaptations to work in your school context, is there an opportunity to develop and evaluate this with a researcher? Have you observed the use of this approach/intervention to judge the quality of implementation? Consider using a quality assurance checklist to make sure it is being used to its potential.

DISCUSSION

Gorard (2020) highlighted that some schools are adopting educational provision with haphazard and disjointed evidence. In the current paper, we have presented an adaptation of an established framework that allows us to think systematically about evidence generation, research development, and translation into educational practice. By considering different types of research (e.g., field testing, efficacy, effectiveness, maintenance) and different methods (e.g., RCTs, small *N* designs, focus groups), we can begin to develop a more holistic understanding of the capabilities of an approach/intervention across different contexts and scales.

In this paper, we argue for researchers to adopt a *question-to-methods* approach, and that research questions should drive the type of methods that we use in education research. Adopting a mixed-methods approach to research can help us address questions relating to the efficacy and effectiveness of an intervention in education, as well as providing us with insight into factors that affect accurate and sustained implementation. Moreover, the use of multiple approaches to address the same research question (i.e., triangulation) can help us to reap the benefits of different designs and compensate for the limitations inherent in all research methods (Shenton, 2004).

Research in education should be an interactive and iterative process, with field testing and refinement at its core. Beyond this, researchers can investigate the impact of an intervention or teaching approach across diverse learner and classroom contexts, as well as across different scales and modalities of implementation. As outlined here, we also need to consider the factors that affect the translation of research into practice, including identifying barriers that make implementation challenging. This may lead to further refinement of an intervention or approach to allow for reasonable adjustments, or the evaluation of supplementary professional training opportunities that might be required.

Several papers provide guidance to enable readers to think critically about the design of, and results from, research. This includes Gorard's (2014) sieve to evaluate the

trustworthiness of group-design studies against six domains: design, scale, dropout, outcomes, fidelity and validity. The single case experimental design (SCED) scale allows us to rate the quality of reporting within small N studies, looking for details such as participant demographics, definitions of target behaviour, and inter-rater reliability scores (Tate et al., 2008). Moving towards focus groups, interviews and qualitative surveys, Shenton (2004) suggested that we should consider four criteria: credibility, transferability, dependability and confirmability. Discussion papers such as these serve two main functions:

1. For readers of research to draw their own conclusions about how to interpret the results, and consider its contribution to the evidence-base
2. For researchers to think critically about the design of future studies, to ensure their results are as trustworthy as possible. Researchers should be transparent about the limitations of their research and any associated bias that they may have contributed to the process.

It is important to make the distinction between the terms *evidence-informed* and *evidence-based* in discourse about educational interventions. Here we propose that a new approach/intervention is *evidence-informed* during phases 0 and 1 of this framework. During a review of previous literature, curriculum designers can draw from pedagogic theory about how pupils develop relevant skills. This can inform the development of the instructions, and activities that underlie the new approach/intervention. After the field-testing phase, researchers can begin to collect empirical data to validate its efficacy and effectiveness in a particular context. An accumulation of studies demonstrating positive impact on learner outcomes across phases 2 and 3 contribute to the evolving evidence-base for that approach/intervention. This is a prerequisite to consider an approach/intervention as *evidence-based*.

We should not discredit educational provision that only has data to support its use under efficacious conditions. Sometimes approaches only show desired effects when delivered to high levels of fidelity. If it is the most cost-effective and/or impactful intervention that is available for the given context, then leaders may need to consider out-sourcing support to facilitate its use in this way. This may fit into the Response-to-Intervention (RtI) framework, whereby a small proportion of a class requires intensive one-to-one or small group learning support delivered by specialists (Hughes & Dexter, 2011).

Some interventions can demonstrate desired effects even under sub-optimal implementation conditions (i.e., where fidelity is <100%). For an approach/intervention to be adopted by schools on a widespread basis, it may be more practical for a teacher or teaching assistant to deliver it. However, it is important to reiterate that provision is only *evidence-based* within the boundaries of the research context in which it has been investigated. If a teacher or teaching assistant wants to implement it without specialist support, then it is imperative that they consider if there is promising evidence to support it across phases 3 and 4.

It is also important to conduct research with different learner populations (e.g., mainstream learners and those with additional learning needs), across different scales (1:1, small group, whole-class, whole-school), and with varying levels of fidelity (efficacy versus effectiveness research). The accumulation of this research will allow us to understand if an approach/intervention is likely to be effective when used in everyday educational settings. Thornicroft et al. (2011) outlined that it is important to conceptualise the evidence-building process as a journey consisting of key phases leading to the accumulation of relevant and contextualised evidence rather than an over-reliance on single studies that might not form part of a sequential evidence-building pathway.

Additional considerations

Equity and language

Researchers and educators working in bilingual contexts also need to be aware of the idiosyncrasies of language (Share, 2008). If an intervention has a promising evidence-base through the medium of English, we cannot assume it will be as equally as effective if adapted in a different language (Jones et al., 2011; Mueller Gathercole et al., 2008). Each language comes with its own vocabulary, syntax and grammar affecting the conceptualisation of meaning (Esposito, 2001; Larson, 1998). At phase 0, we need to consider whether specific adjustments need to be made to provision to ensure linguistic fidelity and cultural appropriateness—particularly with reference to interventions centred on language development and communication, or those that rely on scripted delivery. Across phases 2 to 5, researchers and teachers may need to consider the languages of the assessments that they use to evaluate impact to ensure that they are both appropriate and accessible. For example, Jones et al. (2006) explained that few questionnaires are available in American Sign Language, preventing those who communicate using this system from participating in research. It is important that translated materials show high levels of internal reliability with any other language format that has been validated. Doing this will help improve the inclusivity of practice and research, across the curriculum and lifespan of education.

Caveats when reviewing research

The process of research generation can be slow and onerous. This is because larger trials often require researchers to bid for competitive grant funding. A funded trial then requires researchers to recruit schools and pupils to participate, the intervention period requires time to run with assessments on either side, and then researchers need to analyse and report the outcomes. Beyond this, the process to publish research often requires iterations of peer-review and revisions. It can take a considerable amount of time for research around effective strategies and practice to become available to schools, policy makers, and other researchers. This also means that summaries of evidence available in repositories such as the Education Endowment Foundation do not necessarily capture the most recent developments in the field.

Not all useful research gets published or disseminated to educators. There is a bias in education to publish studies that include large samples and demonstrate educationally significant results. This is problematic because it means that systematic reviews and meta-analyses cannot accurately reflect all the quality research that relates to their research questions and limits the ability of policy makers to make fully informed decisions (Banks et al., 2012). Moreover, it means other researchers may waste resources conducting research that has already been done or focus on interventions that are ineffective. We need to be mindful of this in our decision making and consider other outlets for findings that do not make it through traditional publication.

As researchers, we have an ethical responsibility to make the results of our research more widely available through accessible formats (British Educational Research Association, 2018b). To help bridge the research-to-practice gap, researchers need to produce more accessible summaries of research for schools and other key stakeholders. These summaries should outline the key findings from more trustworthy studies, including limitations and practical implications for educators.

The need for cost-effectiveness research

Levin and Belfield (2015) argued that cost-effectiveness analysis is underused in education research. Decision makers who are responsible for resource allocation are often unaware of how cost-effectiveness data can enhance practice in schools. For both policy makers and teachers, grappling with this evidence can help increase the efficiency of resource use within classrooms. By selecting interventions with the highest evidence of effectiveness relative to their cost, we could improve productivity in education (Hollands et al., 2016). As researchers, we should report sufficient detail about the design of our studies, such as the cost of tangible resources, dosage (i.e., the number of hours each week pupils use the intervention), training, scale, assessments, details about follow-on support, and the costs incurred by schools to run the intervention. Cost-effectiveness evaluations enable school leaders to make better informed decisions about which provision is best for their learners and budget.

Comparing cost-effectiveness across studies is hindered by the disparity of the skills assessed in impact measures (Kraft, 2020). For studies that use the same measures to assess the impact of the same intervention, we can compare findings across different populations, scales and possibly even language contexts. When reviewing studies that assess similar provision using different outcome measures, we may need to cast a more critical eye over the nuances of programme design to establish which is most appropriate for the contexts in which we work, relative to cost. Once we have established if an approach/intervention has promise in a particular context, we can begin to compare it with alternatives to help schools make informed decisions on how to best invest time and money, whilst providing the most benefit to pupils (Willingham & Daniel, 2021). To support this, researchers and policy makers should take a coordinated approach towards the generation, and open-access curation, of cost-effectiveness evidence.

A note for policy makers

We should be wary of the widespread adoption of programmes that lack robust evaluations and/or cannot be shown to work in everyday school settings (Pegram et al., 2022). Over the years, research has found several widely adopted educational theories such as learning styles (Evidence for Learning, n.d.; Pashler et al., 2009) and brain training apps (Spaulding et al., 2010; Stephenson, 2009) to have little positive impact on learner outcomes and/or be theoretically flawed. The continued investment in under-developed interventions and teaching approaches (particularly within phase 0 and 1) is not a good use of public money. It is imperative for education funders and researchers to evaluate promising approaches within an evidence-generating framework.

Evidence generation can be largely driven by supply-side stakeholders from left to right on Figure 1 (Perkins et al., 2007; Thornicroft et al., 2011), where curriculum designers and/or researchers commonly instigate the creation of new approaches/interventions through phases 0–1. The rest of the journey is contingent on successful funding from educational funding or government bodies. We must ensure that teachers and/or learners also have opportunities to influence the demand-side of evidence generation within this framework so that the development of new approaches and programmes adequately address the needs of stakeholders. This is an aspect that should be addressed by researchers and funders via a more coordinated approach.

In terms of investing in human capital, it is important to invest money wisely into educational provision that is effective. This provision may vary depending on the institution making use of it (e.g., mainstream schools, special education schools, alternative provision) and

indeed for pupils' varying needs. To identify cost-effective provision that has the potential to scale to tier 1 or 2 within the Rtl framework, we need to ensure we have the evidence to suggest that the intervention can lead to positive effects within phase 3 and challenges associated within phase 4 have been mitigated (e.g., through the offering of meaningful professional development opportunities to support implementation).

Moving towards a coordinated approach

We believe that researchers, policy makers, funders, and other key stakeholders should coordinate efforts to ensure that:

- Proposals for new research and/or new education provision commence with a thorough review of the existing evidence (phase 0). This involves casting a critical eye over the trustworthiness of existing research and the type of question it answers within phases 1–4. The development of new interventions should be grounded in logical theory and other appropriate basic science.
- We develop logic models to support the systematic generation of evidence and consider the longevity of our efforts. Logic models can be useful to present the detail for individual studies and a wider scheme of research work (to see how the findings complement and build on each other). Where necessary, these should be altered to reflect the iterations and refinements to the intervention.
- Any findings about barriers that impact the outcomes and/or prevent translation between research phases should be retained and reported through an open access channel to help contribute to knowledge dissemination.
- Economic costs associated with each research study are reported to enable for subsequent cost-effectiveness analyses. This will help ensure that schools can make informed decisions about what to invest in after the termination of a research study and enable us to make relative comparisons across similar interventions.
- High-quality evidence is disseminated in an accessible format via a respected conduit.
- Sufficient training opportunities are available to support the use of the approach/intervention at school. Depending on the available evidence (efficacy vs. effectiveness), we need to consider whether specialists or educators are the best placed to deliver the approach/intervention.
- We support reform efforts to help teachers as they realise curricula in ways that are appropriate for the learners in their classroom (e.g., Furlong et al., 2021). This includes providing support to help develop a research literate workforce who feel confident to engage in professional enquiry to improve their practice (Kelchtermans, 2021). Griffiths et al. (2020) discuss how this training can be integrated into initial teacher education (ITE) courses.

CONCLUSION

We have shown how researchers and practitioners can use an evidence building framework (adapted from medical and psychosocial science) to conceptualise the journey from evidence-informed to evidence-based for educational provision. New approaches/interventions require evidence across several phases, from basic science through to implementation, and we can face several challenges along this journey that hinder knowledge creation and mobilisation. This framework provides a way for researchers and research-informed professionals to critically appraise existing evidence, identify questions that still need answering, and plan for long-term research programmes that build on knowledge. We also argue that this framework encourages us to think systematically about which questions need answering next and the most appropriate research methods to answer these (i.e., follow a *questions-to-methods*

approach). This naturally leads to a mixed methods approach and a recognition that multiple methods are required to build a holistic and strong evidence-base for educational provision. We also distinguish between enquiry and research and argue for a model of *enquiry-to-research* that helps us understand the different roles required by stakeholders in building and implementing evidence, and when and how they should collaborate. Throughout this paper we have intentionally included a range of key questions and a more accessible summary of each phase in the framework to help practitioners engage with this journey.

At a system level, the generation of relevant and trustworthy evidence can help school leaders and practitioners make better informed decisions about how they invest resources (including time and money) for maximum impact on learner outcomes. As we work towards embedding evidence-based practice in schools, it is also important to acknowledge some of the challenges that prevent the uptake and maintenance of promising approaches. This includes looking at the translational barriers that hinder the application of findings from one phase to the next and listening to those who have used the approach. Once we acknowledge these barriers, we can think systematically about how we might mitigate and overcome them. Our aim has been to contribute to the debate in education about how we might identify and develop evidence-based approaches that have the strongest utility, and how we can foster fruitful collaborations between key stakeholders.

CONFLICT OF INTEREST

No conflicts to disclose.

ETHICAL APPROVAL

Ethical approval was not required for this article. We did not collect or analyse any data as part of this review/discussion piece.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the manuscript.

ORCID

Kaydee L. Owen  <https://orcid.org/0000-0002-8198-1175>

Richard C. Watkins  <https://orcid.org/0000-0001-9584-0747>

J. Carl Hughes  <https://orcid.org/0000-0001-7130-7286>

REFERENCES

- Alika, I. J., & Aibieyi, S. (2014). Human capital: Definitions, approaches and management dynamics. *Journal of Business Administration and Education*, 5(1), 55–78.
- Arthur, M. W., & Blitz, C. (2000). Bridging the gap between science and practice in drug abuse prevention through needs assessment and strategic community planning. *Journal of Community Psychology*, 28(3), 241–255.
- Banks, G. C., Kepes, S., & Banks, K. P. (2012). Publications bias: The antagonist of meta-analytic reviews and effective policymaking. *Educational Evaluation and Policy Analysis*, 34(4), 259–277. <https://doi.org/10.3102/0162373712446144>
- Barger-Anderson, R., Domaracki, J. W., Kearney-Vakulick, N., & Kubina, R. M. (2004). Multiple baseline designs: The use of a single-case experimental design in literacy research. *Reading Improvement*, 41(4), 217–255.
- Barnett, S. D., Heinemann, A. W., Libin, A., Houts, A. C., Gassaway, J., Sen-Gupta, S., Resch, A., & Brossart, D. F. (2012). Small N designs for rehabilitation research. *Journal of Rehabilitation Research and Development*, 49(1), 175–186. <https://doi.org/10.1682/JRRD.2010.12.0242>
- Barrio Minton, C. A., Wachter Morris, C. A., & Yaites, L. D. (2014). Pedagogy in counselor education: A 10-year content analysis of journals. *Counselor Education and Supervision*, 53(3), 162–177. <https://doi.org/10.1002/j.1556-6978.2014.00055.x>
- Berger, N., & Fisher, P. (2013). *A well-educated workforce is key to state prosperity*. Economic Policy Institute. Retrieved from <https://www.epi.org/publication/states-education-productivity-growth-foundations/>

- Binder, C. (2003). Doesn't everybody need fluency? *Performance Improvement*, 42(3), 14–20. <https://doi.org/10.1002/pfi.4930420304>
- Boyle, J., & Kelly, B. (2017). The role of evidence in educational psychology. In B. Kelly, L. Woolfson, & J. Boyle (Eds.), *Frameworks for practice in educational psychology: A textbook for trainees and practitioners* (2nd ed., pp. 29–43). Jessica Kingsley.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage.
- British Educational Research Association (2018). *The BERA close-to-practice research report*. Retrieved from https://www.lse.org.uk/uploads/datahub/4567ceb%5E03in-03/2018-12-10-BERA-Close-to-Practice_report_Nov2018.pdf
- British Educational Research Association (2018). *Ethical guidelines for educational research, fourth edition*. Retrieved from <https://www.bera.ac.uk/publication/ethical-guidelines-for-educational-research-2018>
- Coldwell, M., & Maxwell, B. (2018). Using evidence-informed logic models to bridge methods in educational evaluation. *Review of Education*, 6(3), 267–300. <https://doi.org/10.1002/rev3.3151>
- Connolly, P., Biggart, A., Miller, S., O'Hare, L., & Thurston, A. (2017). *Using randomised controlled trials in education*. Sage.
- Department for Education (2022). *Guidance: Choosing a phonics teaching programme*. Retrieved from <https://www.gov.uk/government/publications/choosing-a-phonics-teaching-programme/list-of-phonics-teaching-programmes>
- Detrich, R. (2020). Cost-effectiveness analysis: A component of evidence-based education. *School Psychology Review*, 49(4), 423–430. <https://doi.org/10.1080/2372966X.2020.1827864>
- Drew, V., Priestley, M., & Michael, M. K. (2016). Curriculum development through critical collaborative professional enquiry. *Journal of Professional Capital and Community*, 1(1), 92–106. <https://doi.org/10.1108/JPC-09-2015-0006>
- Eagleton, S. (2017). Designing blended learning interventions for the 21st century student. *Advanced Physiological Education*, 41, 203–211. <https://doi.org/10.1152/advan.00149.2016>
- Edoald, T., & Nevill, C. (2021). Working out what works: The case of the Education Endowment Foundation in England. *ECNU Review of Education*, 4(1), 46–64. <https://doi.org/10.1177/2096531120913039>
- Education Endowment Foundation (2019). *Putting evidence to work: A school's guide to implementation, guidance report*. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Publications/Implementation/EEF_Implementation_Guidance_Report_2019.pdf
- Education Endowment Foundation (2021). *Effective professional development: Guidance report*. Retrieved from <https://d2tic4wvo1iusb.cloudfront.net/guidance-reports/effective-professional-development/eef-effective-pd-report.pdf>
- Egan, D. (2016). Educational equity in Wales. *Wales Journal of Education*, 18(1), 21–44. <https://doi.org/10.16922/wje.18.1.4>
- Egan, T. E., Wymbs, F. A., Owens, J. S., Evans, S. W., Hutus, C., & Allan, D. M. (2019). Elementary school teachers' preference for school-based interventions for students with emotional and behavioral problems. *Psychology in the Schools*, 56(10), 1633–1653. <https://doi.org/10.1002/pits.22294>
- Elliot, J. (1991). *Action research for educational change*. McGraw-Hill Education.
- Engelmann, S., & Carnine, D. W. (1982). *Theory of instruction: Principles and applications*. Irvington.
- Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2), 43–71. <https://doi.org/10.1002/piq.21143>
- Esposito, N. (2001). From meaning to meaning: The influence of translation techniques on non-English focus group research. *Qualitative Health Research*, 11(4), 568–579. <https://doi.org/10.1177/104973201129119217>
- Evidence for Learning (n.d.). *Learning styles*. Retrieved from <https://evidenceforlearning.org.au/the-toolkits/the-teaching-and-learning-toolkit/all-approaches/learning-styles/#how-effective-is-it>
- Flannery, K. B., Sugai, G., & Anderson, C. M. (2009). School-wide positive behavior support in high school: Early lessons learned. *Journal of Positive Behavior Interventions*, 11(3), 177–185. <https://doi.org/10.1177/1098300708316257>
- Fletcher, S., & Mullen, C. A. (Eds.). (2012). *SAGE handbook of mentoring and coaching in education*. SAGE.
- Furlong, J., Griffiths, J., Hannigan-Davies, C., Harris, A., & Jones, M. (2021). The reform of initial teacher education in Wales: From vision to reality. *Oxford Review of Education*, 47(1), 61–78. <https://doi.org/10.1080/03054985.2020.1842180>
- Gibson, J. E. (2012). Interviews and focus groups with children: Methods that match children's developing competencies. *Journal of Family Theory & Review*, 4, 148–159. <https://doi.org/10.1111/j.1756-2589.2012.00119.x>
- Gorard, S. (2014). A proposal for judging the trustworthiness of research findings. *Radical Statistics*, 110, 47–59.
- Gorard, S. (2020). *Getting evidence into education: Evaluating the routes to policy and practice*. Routledge.
- Graham, J. E., Karmarkar, A. M., & Ottenbacher, K. J. (2012). Small sample research designs for evidence-based rehabilitation: Issues and methods. *Archives of Physical Medicine and Rehabilitation*, 93(8), 111–116. <https://doi.org/10.1016/j.apmr.2011.12.017>

- Griffiths, J., Bamber, S., French, G., Hulse, B., Jones, G., Jones, R. C., Jones, S., Williams, G. M., Wordsworth, H., & Hughes, J. C. (2020). Growing tomorrow's teachers together: The CaBan initial teacher education partnership. *Wales Journal of Education*, 22(1), 209–231. <https://doi.org/10.16922/wje.22.1.10-en>
- Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6), 2052–2086. <https://doi.org/10.1257/aer.103.6.2052>
- Higgins, S. (2020). The development and worldwide impact of teaching and learning toolkit. In S. Higgins (Ed.), *Getting evidence into education: Evaluating the routes to policy and practice* (pp. 69–83). Routledge.
- Hollands, F. M., Kieffer, M. J., Shand, R., Pan, Y., Cheng, H., & Levin, H. M. (2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. *Journal of Research on Educational Effectiveness*, 9(1), 30–53. <https://doi.org/10.1080/19345747.2015.1055639>
- Hollands, F., Pan, Y., & Escueta, M. (2019). What is the potential for applying cost-utility analysis to facilitate evidence-based decision making in schools? *Educational Researcher*, 48(5), 287–295. <https://doi.org/10.3102/0013189X19852101>
- Hughes, C. A., & Dexter, D. D. (2011). Response to intervention: A research-based summary. *Theory into Practice*, 50(1), 4–11. <https://doi.org/10.1080/00405841.2011.534909>
- Huitt, W., Huitt, M., Monetti, D., & Hummel, J. (2009). *A systems-based synthesis of research related to improving students' academic performance*. Paper presented at the 3rd International City Break Conference sponsored by the Athens Institute for Education and Research (ATINER), October 16–19, Athens, Greece. Retrieved from <http://www.edpsycinteractive.org/papers/improving-school-achievement.pdf>
- Jones, E. W., Hoerger, M., Hughes, J. C., Williams, B. M., Jones, B., Moseley, Y., Hughes, D. R., & Prys, D. (2011). ABA and diverse cultural and linguistic environments: A Welsh perspective. *Journal of Behavioral Education*, 20(4), 297–305. <https://doi.org/10.1007/s10864-011-9138-5>
- Jones, E. G., Mallinson, R. K., Phillips, L., & Kang, Y. (2006). Challenges in language, culture, and modality. *Nursing Research*, 55(2), 75–81.
- Keeley, B. (2007). *Human capital: How what you know shapes your life*. OECD: OECD Insights.
- Kelchtermans, G. (2021). Keeping educational research close to practice. *British Educational Research Journal*, 47(6), 1504–1511. <https://doi.org/10.1002/berj.3716>
- Kortessluoma, R. L., Hentinen, M., & Nikkonen, M. (2003). Conducting a qualitative child interview: Methodological considerations. *Journal of Advanced Nursing*, 42(5), 434–441. <https://doi.org/10.1046/j.1365-2648.2003.02643.x>
- Koziuff, M. A., LaNunziata, L., Cowardin, J., & Bessellieu, F. B. (2000). Direct instruction: Its contributions to high school achievement. *The High School Journal*, 84(2), 54–71.
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Kretlow, A. G., & Bartholomew, C. C. (2010). Using coaching to improve the fidelity of evidence-based practices: A review of studies. *Teacher Education and Special Education*, 33(4), 279–299. <https://doi.org/10.1177/0888406410371643>
- Krnic Martinic, M., Pieper, D., Glatt, A., & Puljak, L. (2019). Definition of a systematic review used in overviews of systematic reviews, meta-epidemiological studies and textbooks. *BMC Medical Research Methodology*, 19(1), 1–12. <https://doi.org/10.1186/s12874-019-0855-0>
- Kubina, R. M., Morrison, R., & Lee, D. L. (2002). Benefits of adding precision teaching to behavioural interventions for students with autism. *Behavioral Interventions*, 17, 233–246. <https://doi.org/10.1002/bin.122>
- Lanovaz, M. J., Turgeon, S., Cardinal, P., & Wheatley, T. L. (2019). Using single-case designs in practical settings: Is within-subject replication always necessary? *Perspectives on Behavior Science*, 42, 153–162. <https://doi.org/10.1007/s406-018-0138-9>
- Larson, M. L. (1998). *Meaning-based translation: A guide to cross-language equivalence* (2nd ed.). University Press of America.
- Law, J., Dennis, J. A., & Charlton, J. J. (2017). Speech and language therapy interventions for children with primary speech and/or language disorders. *The Cochrane Database of Systematic Reviews*, 2017(1). <https://doi.org/10.1002/14651858.CD012490>
- Layng, T. V. (2018). Tutorial: Understanding concepts: Implications for behavior analysts and educators. *Perspectives on Behavior Science*, 42, 345–363. <https://doi.org/10.1007/s40614-018-00188-6>
- Lee, O., Penfield, R., & Maerten-Rivera, J. (2009). Effects of fidelity of implementation on science achievement gains among English language learners. *Journal of Research in Science Teaching*, 46(7), 836–859. <https://doi.org/10.1002/tea.20335>
- Levin, H. M., & Belfield, C. (2015). Guiding the development and use of cost-effectiveness analysis in education. *Journal of Research on Educational Effectiveness*, 8(3), 400–418. <https://doi.org/10.1080/19345747.2014.915604>

- MacLean, M. S., & Mohr, M. M. (1999). *Teacher-Researchers at work*. National Writing Project.
- Meyer, M. M., & Behar-Horenstein, L. S. (2015). When leadership matters: Perspectives from a teacher team implementing response to intervention. *Education and Treatment of Children, 38*(3), 383–402. <https://doi.org/10.1353/etc.2015.0022>
- Mueller Gathercole, V. C., Thomas, E. M., & Hughes, E. (2008). Designing a normed receptive vocabulary test for bilingual populations: A model from Welsh. *International Journal of Bilingual Education and Bilingualism, 11*(6), 678–720. <https://doi.org/10.1080/13670050802149283>
- O'Reilly, M., Svirydzhenka, N., Adams, S., & Dogra, N. (2018). Review of mental health promotion interventions in schools. *Social Psychiatry and Psychiatric Epidemiology, 53*, 647–662. <https://doi.org/10.1007/s00127-018-1530-1>
- Owen, K. L., Hunter, S. H., Watkins, R. C., Payne, J. S., Bailey, T., Gray, C., Hastings, R. P., & Hughes, J. C. (2021). Implementation support improves outcomes of a fluency-based mathematics strategy: A cluster-randomized controlled trial. *Journal of Research on Educational Effectiveness, 14*(3), 523–542. <https://doi.org/10.1080/19345747.2021.1875526>
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2009). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119. <https://doi.org/10.1111/j.1539-6053.2009.01038.x>
- Pegram, J., Watkins, R. C., Hoerger, M., & Hughes, J. C. (2022). Assessing the range and evidence-base of interventions in a typical school cluster. *Review of Education*. <https://doi.org/10.1002/rev3.3336>
- Perkins, M. B., Jensen, P. S., Jaccard, J., Gollwitzer, P., Oettingen, G., Pappadopulos, E., & Hoagwood, K. E. (2007). Applying theory-driven approaches to understanding and modifying clinicians' behavior: What do we know? *Psychiatric Services, 58*(3), 342–348. <https://doi.org/10.1176/ps.2007.58.3.342>
- Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research, 90*(1), 24–46. <https://doi.org/10.3102/0034654319877153>
- Polanin, J. R., Tanner-Smith, E. E., & Hennessy, E. A. (2016). Estimating the difference between published and unpublished effect sizes: A meta-review. *Review of Educational Research, 86*(1), 207–236. <https://doi.org/10.3102/0034654315582067>
- Reich, N. G., & Milstone, A. M. (2014). Improving efficiency in cluster-randomized study design and implementation: Taking advantage of a crossover. *Open Access Journal of Clinical Trials, 6*, 11–15. <https://doi.org/10.2147/OAJCT.S56730>
- Rosenshine, B. (2012). Principles of Instruction: research-based strategies that all teachers should know. *American Educator, 36*, 12–39. Retrieved from <https://www.teachertoolkit.co.uk/wp-content/uploads/2018/10/Principles-of-Instruction-Rosenshine.pdf>
- Sailors, M., & Shanklin, N. L. (2010). Introduction: Growing evidence to support coaching in literacy and mathematics. *The Elementary School Journal, 111*(1), 1–6. <https://doi.org/10.1086/653467>
- Schön, D. (1983). *The reflective practitioner*. Basic Books.
- Shand, R., & Bowden, A. B. (2021). Empirical support for establishing common assumptions in cost research in education. *Journal of Research on Educational Effectiveness*. Advance online publication. <https://doi.org/10.1080/19345747.2021.1938315>
- Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584–615. <https://doi.org/10.1037/0033-2909.134.4.584>
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75. <https://doi.org/10.3233/EFI-2004-22201>
- Sims, S., & Fletcher-Wood, H. (2021). Identifying the characteristics of effective professional development: A critical review. *School Effectiveness and School Improvement, 32*(1), 47–63. <https://doi.org/10.1080/09243453.2020.1772841>
- Slavin, R. E. (2020). How evidence-based reform will transform research and practice in education. *Educational Psychologist, 55*(1), 21–31. <https://doi.org/10.1080/00461520.2019.1611432>
- Slocum, T. A., Detrich, R., Wilczynski, S. M., Spencer, T. D., Lewis, T., & Wolfe, K. (2014). The evidence-based practice of applied behavior analysis. *The Behavior Analyst, 37*(1), 41–56. <https://doi.org/10.1007/s40614-014-0005-2>
- Spaulding, L. S., Mostert, M. P., & Beam, A. P. (2010). Is Brain Gym® an effective educational intervention? *Exceptionality, 18*(1), 18–30. <https://doi.org/10.1080/09362830903462508>
- Stephenson, J. (2009). Best practice? Advice provided to teachers about the use of Brain Gym® in Australian schools. *Australian Journal of Education, 53*(2), 109–124. <https://doi.org/10.1177/000494410905300202>
- Stoa, R., & Chu, A. (2020). An argument for implementing and testing novelty in the classroom. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. <https://doi.org/10.1037/stl000023>
- Tate, R. L., McDonald, S., Perdices, M., Togher, L., Schultz, R., & Savage, S. (2008). Rating the methodological quality of single-subject designs and n-of-1 trials: Introducing the Single-Case Experimental Design (SCED) Scale. *Neuropsychological Rehabilitation, 18*(4), 385–401. <https://doi.org/10.1080/09602010802009201>

- Taylor-Powell, E., & Henert, E. (2008). *Developing a logic model: Teaching and training guide*. University of Wisconsin-Extension, Cooperative Extension, Program Development and Evaluation. Retrieved from <https://fyi.extension.wisc.edu/programdevelopment/files/2016/03/lmguidecomplete.pdf>
- Thomas, G. (2020). Experiment's persistent failure in education inquiry, and why it keeps failing. *British Educational Research Journal*, 47(3), 501–519. <https://doi.org/10.1002/berj.3360>
- Thornicroft, G., Lempp, H., & Tansella, M. (2011). The place of implementation science in the translational medicine continuum. *Psychological Medicine*, 41(11), 2015–2021. <https://doi.org/10.1017/S0033291711000109>
- Tyler, E. J., Watkins, R. C., Roberts, S. E., Hoerger, M., Hastings, R. P., Hulson-Jones, A., & Hughes, C. J. (2019). The collaborative institute for education research, evidence and impact: A case study in developing regional research capacity in Wales. *Wales Journal of Education*, 21(1), 89–108. <https://doi.org/10.16922/wje.21.1.6>
- White, H. (2019). The twenty-first century experimenting society: The four waves of the evidence revolution. *Palgrave Communications*, 5, 1–7. <https://doi.org/10.1057/s41599-019-0253-6>
- Willingham, D. T., & Daniel, D. B. (2021). Making education research relevant: How researchers can give teachers more choices. *Education Next*, 21(2), 28–34.
- Wrigley, T. (2018). The power of 'evidence': Reliable science or a set of blunt tools? *British Educational Research Journal*, 44(3), 359–376. <https://doi.org/10.1002/berj.3338>
- Wyse, D., Brown, C., Oliver, S., & Poblete, X. (2018). *The BERA close-to-practice research project: Research report*. Retrieved from <https://www.bera.ac.uk/publication/bera-statement-on-close-to-practice-research>
- Wyse, D., Brown, C., Oliver, S., & Poblete, X. (2021). Education research and educational practice: The qualities of a close relationship. *British Educational Research Journal*, 47(6), 1466–1489. <https://doi.org/10.1002/berj.3626>

How to cite this article: Owen, K. L., Watkins, R. C., & Hughes, J. C. (2022). From evidence-informed to evidence-based: An evidence building framework for education. *Review of Education*, 10, e3342. <https://doi.org/10.1002/rev3.3342>