Research Paper

ROE

ECNU Review of Education
2021, Vol. 4(1) 7–24

DOI: 10.1177/2096531120976060
journals.sagepub.com/home/roe

SAGE

ECNUP

How Could Evidence-Based Reform Advance Education?

Robert E. Slavin
Johns Hopkins University

Alan C. K. Cheung
The Chinese University of Hong Kong

Tengteng Zhuang（庄腾腾）
Beijing Normal University

Abstract

Purpose: This article presents a definition and rationale for evidence-based reform in education, and a discussion of the current state of evidence-based research, focusing on China, the U.S., and the UK.

Design/Approach/Methods: The article contrasts the state of educational research in the U.S., the UK, and China, world leaders in evidence-based reform.

Findings: The article suggests ways in which Chinese, U.S., UK, and other scholars might improve the worldwide quality of evidence-based reform in education. One indicator of this partnership is an agreement among the Chinese University of Hong Kong, Nanjing Normal University, and Johns Hopkins University to work together on Chinese and English versions of the website Best Evidence in Brief and a collaboration between Johns Hopkins and the *ECNU Review of Education* at East China Normal University.

Originality/Value: This article is the first to compare developments in evidence-based reform in education in China with parallel developments in the U.S. and UK Building understanding of

Corresponding author:
Robert E. Slavin, Center for Research and Reform in Education, Johns Hopkins University, Baltimore, MD 21218, USA.
Email: rslavin@jhu.edu

current evidence-based policies should help all countries learn how evidence can play a greater role in education policy and benefit many students.

Keywords

Evidence-based reform in education, experiments, international comparisons, meta-analysis, research designs

Date received: 16 March 2020; accepted: 27 May 2020

Evidence-based reform in education (Bridgeland & Orszag, 2013; Gueron & Rolston, 2013; Slavin, 2008, 2013, 2017, 2020; Slavin & Cheung, 2017) refers to policies in which educators and policymakers use evidence of effectiveness as a criterion for choosing educational programs, products, and practices. Evidence of effectiveness is defined as evidence from rigorous experiments in which students experiencing experimental programs are compared over significant periods of time (say a semester or more) to those using traditional control methods in terms of gains on valid measures of achievement. Ideally, students, teachers, and/or schools are assigned at random to experimental or control treatments (Bucks & McGee, 2015; Gueron & Rolston, 2013), but at a minimum, experimental and control students are well matched at pretest on measures of achievement, demographic variables, and so on. Note that this definition does not refer to simply disseminating information about generic principles of effective practice, which has not generally been found to make much of a difference in practices or outcomes (e.g., Hemsley-Brown & Sharp, 2003; Weiss et al., 2008), but instead refers to reforms that support the use of specific programs evaluated in comparison to control groups and found to be effective and replicable.

Although all types of research methodologies are used in evidence-based reform, the distinctive tool for determining outcomes of innovative approaches compared to ordinary practices is the experiment, or quasi-experiment. In experiments in education, schools, classes/teachers, or students are randomly assigned to treatment groups, which implement the innovative methods, or to control groups, which continue usual practices. Quasi-experiments are similar, but instead of random assignment, quasi-experiments match schools, classes/teachers, or students based on prior achievement and other factors. In both cases, the level of clustering is important. Randomly assigning or matching schools is considered ideal, because there is always a concern about contamination when random assignment or matching takes place within schools. For example, control teachers may hear about the new methods and start using them in their own classes. However, when clustering is used at the school or class/teacher levels, analyses must take clustering into account, so studies with school-level clustering usually need 40–50 units (schools or classes). This is not the place for a detailed discussion of methods, but it is important to know that most experiments and quasi-experiments are quite large, due to their use of clustering.

Why should evidence matter?

Why should evidence of effectiveness be a major criterion in the selection of educational products and services? The most important answer is that programs with a strong evidence base that are implemented as they were in the validating research are likely to produce better outcomes for students. Further, making evidence a basis for program adoption would put education into a virtuous cycle of innovation, evaluation, and progressive improvement like that which has transformed fields such as medicine, agriculture, and technology (Haskins, 2014; Kolada, 2013).

Evidence-based reform is, at its core, a simple idea. It means engaging in four essential activities (see Slavin, 2013):

1. Finding out what works best in every subject and grade level of education, using well-established experimental methods to evaluate innovative programs.
2. Making educators and policymakers aware of proven programs and practices.
3. Providing resources and incentives to enable school leaders to implement proven programs and practices.
4. Creating policies and systems to continually add to knowledge and effective practice.

These four principles underlie policies and practices in many areas known for rapid and sustained progress on a large scale. For example, the astonishing advances in medicine and public health over the past century are due to evidence-based medicine, where governments have supported research and development of new medicines, techniques, equipment, and public health policies, have sponsored rigorous evaluations of these ideas, and have proactively disseminated treatments and policies known to be effective. As a result, national average life expectancy has increased dramatically in China and other countries, diseases have been cured, and preventive practices, such as vaccines against smallpox, have been institutionalized.

In agriculture, similar transformations have taken place. In China, advances in productivity of rice farming (a tripling of productivity per hectare over the past 50 years) have transformed life, ending the famines that were once common. These and many other advances in agriculture came about as scientists developed and tested new seeds, fertilizer, and insect control as well as improved breeds of animals, veterinary practices, feed, and much more. As effective methods have been proven, they have been disseminated to farmers, making farming more efficient and profitable as well as feeding whole populations. Advances in technology, engineering, and many other successful fields follow the same pattern: research, development, evaluation, and dissemination.

Until recently, education did not participate to any major degree in evidence-based practices or policies. Yet there is no reason education could not follow the lead of medicine, agriculture,

technology, and engineering and begin to base its policies and practice on rigorous evidence. One could argue that education is not the same as medicine or agriculture, because a pill or a new breed of cattle can readily be replicated, while educational innovations always involve many people changing their behaviors in diverse contexts. Yet these conditions also arise in medicine and agriculture, as when physicians learn new ways to maximize compliance with needed lifestyle changes or with taking medications. Farmers may improve yields by using new forms of plowing or management of their water resources. Education has its own evidence base that has shown replicated positive effects of educational methods, such as tutoring and cooperative learning, and there is no reason that these and other research-proven methods could not have a profound impact on educational outcomes.

Evidence-based reform is beginning to transform educational research and practice in some parts of the world. In recent years, Chinese researchers have begun to emphasize the use of rigorous evidence to evaluate educational programs and practices. This article begins with a discussion of evidence-based reform in China, followed by a discussion on evidence-based reform in the U.S. and UK, and concluds with a discussion of how collaboration between Chinese and Western researchers and policymakers could benefit education anywhere in the world where governments encourage evidence-based reform.

Education-based reform in Chinese education

Evidence-based educational research has been a keyword in Chinese academia over the past three decades. Inputting the Chinese phrase“教育实证研究”(educational evidence-based research) in China's largest journal article database, the China National Knowledge Index, yields around 5,000 articles whose titles contain this phrase. There is no doubt that evidence-based educational research has been gaining momentum in Chinese academia in the past two decades. For instance, from 1990 to 2002, there were merely a handful of evidence-based educational studies published each year, with the year 1990 witnessing only two publications of this kind and 2002 yielding 24. However, there has been a dramatic increase in the number of evidence-based research studies since 2003, indicating that the idea of evidence-based research has gradually taken root in Chinese research in education. The volume of evidence-based educational research publications is shown in Table 1.

Educational researchers targeting different levels of education have employed an evidence-based approach to research in China, ranging from compulsory education to vocational education to higher education. The themes and topics studied in an evidence-based approach have been extensive as well, including student satisfaction, educational equity, educational finance, educational expenditure, and even political and ideological education.

Aside from the soaring number of publications, Chinese scholars have placed an increasing emphasis on discussions of various issues pertaining to evidence-based educational research over

Table 1. Number of publications of evidence-based educational research in Chinese academia by year.

Year	Number of publications
1990	2
1992	2
1993	4
1994	3
1995	4
1996	2
1997	2
1998	11
1999	10
2000	14
2001	14
2002	24
2003	45
2004	53
2005	90
2006	144
2007	184
2008	237
2009	244
2010	325
2011	326
2012	352
2013	379
2014	381
2015	387
2016	439
2017	451
2018	477
2019[a]	346

Source. https://www.cnki.net/.

[a]Statistics for the year 2019 are incomplete as of this writing.

the past few years. For instance, East China Normal University (ECNU), a leading higher education institution in educational science in China, has held forums on evidence-based research for five consecutive years since 2015. Its most recent fifth forum (2019) convened scholars nationwide and invited internationally renowned experts. A total of 16 Chinese universities joined ECNU to

Table 2. Themes of parallel sessions of the fifth forum on evidence-based educational research held at East China Normal University in 2019.

No.	Theme	Organizer
1	Educational research based on video-logical analysis	East China Normal University
2	Theoretical reflection and practical review of qualitative research	Beijing Normal University
3	Educational research based on empirical investigation	Northeast China Normal University
4	Educational research based on mixed-methods	Central China Normal University
5	Educational research based on social network analysis	Shaanxi Normal University
6	Innovation in educational discourses based on empirical studies	Southwest University
7	Theoretical discussions on evidence-based educational research	Peking University
8	Experimental research in education	Zhejiang University
9	Action research in education	Xiamen University
10	Theoretical reflection and practical review of quantitative research	Nanjing Normal University
11	Field research in education	Capital Normal University
12	Educational research based on questionnaire design	South China Normal University
13	Case studies in educational research	Zhejiang Normal University
14	Educational research based on ethnography	Northwest Normal University
15	Educational research based on big data analysis	Qufu Normal University
16	Educational research based on text analysis	Jiangxi Institute of Educational Sciences

Source. http://feer.ecnu.edu.cn/4694/list.htm.

jointly organize parallel sessions, with each focusing on the application of evidence-based ideas and methods to one particular aspect of educational research, such as field study, case study, text analysis, and video-logical analysis (Table 2). Numerous subareas are embracing an evidence-based approach for their research findings to be more convincing and robust.

In addition, the School of Education Science at Nanjing Normal University (NNU) has joined the global Best-Evidence-in-Brief (BEiB) network with leading higher education institutions, such as Johns Hopkins University in the U.S. and the Chinese University of Hong Kong, in an effort to disseminate evidence-based research findings to Chinese and worldwide frontline educators on a regular basis. Every 2 weeks, school administrators and teachers and even parents receive the most state-of-the-art evidence-based educational research findings conducted across the world through the member universities of the BEiB network to inform their schools' own educational practices and enhance the quality of education. NNU is so far the first institute of educational science in the Chinese mainland to implement such a mechanism that bridges research and teaching, which

enables frontline educators to understand the best educational practices implemented around the world that improve students' learning outcomes (http://www.cnbeb.org.cn:81/en/about).

Chinese universities such as NNU have also been active in organizing international evidence-based research conferences and forums, with the special purpose of enhancing Chinese scholars' awareness of using evidence-based methods to conduct educational research. For instance, for NNU's 2019 International Conference on Ideas and Actions for Educational Excellence, "evidence-based" is a keyword that frequently appeared in the flyer (http://jky.njnu.edu.cn/info/1042/8496.htm). Key areas discussed in the conference included "Forefronts of International Evidence-Based Research," "Methodological Development of Evidence-Based Educational Research," "China's Evidence-Based Educational Research and Reform," "Evidence-Based Educational Reform and Development," and "Evidence-Based School Improvement."

The words of Professor Zhenguo Yuan, the current Vice Director of the Chinese Society of Education and the Director of the Faculty of Education at ECNU, represented Chinese scholars' understanding of the importance and essentials of evidence-based research. His remarks included the following:

> Evidence-based research is a turning point from speculative philosophy to science, a milestone for the deepening of educational research, an underpinning pillar for theoretical innovation, the only approach to truly resolving disputes and reaching consensus, and a reliable method for the establishment of modern think tanks. (Yuan, 2017)

Even with the rapid emergence of evidence-based publications, numerous Chinese scholars still believe that evidence-based research constitutes a very small proportion of the total research output in Chinese educational academia, and many have pointed out that the pace at which evidence-based research is conducted is in fact not as high as that of non-evidence-based research (Tan, 2017; Wang, 2017). According to Wang (2017), evidence-based research, apart from being premised on a solid theoretical basis, should also count on high-quality data collected and validated statistical methodology to test established educational hypotheses and demonstrate relevant causalities and impact between educational input and output. Non-evidence-based research, on the other hand, mainly refers to studies that adopt illustration of a theory, reasoning over the value of an educational topic based on personal perceptions, plain description of an educational practice, or policy in a foreign country as their research methodology. Wang (2017) also finds that among the publications in five renowned Chinese Social Sciences Citation Index (CSSCI) journals, *Journal of Higher Education*, *Tsinghua Journal of Education*, *China Higher Education Research*, *Peking University Education Review*, and *Fudan Education Forum*, evidence-based studies constituted only about 20% of the total articles published between 2012 and 2016 (Table 3). Viewing the total output by the Chinese researchers of educational sciences, Yuan (2017) argued that evidence-based

Table 3. Publication landscape in the five CSSCI journals between 2012 and 2016.

Journal	Total no. of articles	No. of evidence-based articles	%
Journal of Higher Education	1,150	179	15.6
Tsinghua Journal of Education	527	150	28.5
China Higher Education Research	1,497	257	17.2
Peking University Education Review	338	81	24.0
Fudan Education Forum	587	171	29.1

Source. Wang (2017).
Note. CSSCI = Chinese Social Sciences Citation Index.

research only constitutes less than 10% of the total publication output. These proportions may be accurate, and similar proportions probably exist in the U.S. and UK, where evidence-based research is dominant. Evidence-based research is expensive, so counting numbers of studies, including master's studies and doctoral theses, will always show larger numbers for less quantitative studies. However, among funded research in the U.S. and UK, rigorous experimental research is most common.

The above subheads for the areas, the aforementioned measures that have been taken by Chinese academia, and the call made for more evidence-based studies combined signal the strong interest of contemporary Chinese educational researchers in understanding the development, current status, and practices of evidence-based educational research conducted across the globe, relevant difficulties and possible ways of innovation for carrying out evidence-based studies, and how China itself can implement and apply evidence-based ideas in educational research and reform.

As such, a total of 14 universities and 32 CSSCI journals released a joint declaration, called "East China Normal University Declaration" ("The Declaration" hereafter), on strengthening evidence-based educational research and promoting the shift in research paradigms, in January 2017. The Declaration especially identified 10 measures to be taken in the upcoming years to promote the quality of educational research in China, including:

- Strengthening curriculum development and training of young teachers and researchers on conducting evidence-based research
- Improving evidence-based research output in terms of quantity and quality
- Organizing nationwide evidence-based educational research forums
- Editing and publishing international high-quality publications based on evidence-based research

- Publishing high-quality academic Chinese articles that employ evidence-based research approaches to resolving educational problems
- Improving standards of assessing educational research outputs
- Increasing the proportion of evidence-based educational projects
- Publishing an "Outstanding Achievement Award" for evidence-based research
- Issuing annual reports on the progress of evidence-based educational research progress in China
- Further publicizing educational data to the wider public

(*Source*: ECNU News Center, https://news.ecnu.edu.cn/58/98/c1833a88216/page.htm)

Skepticism of evidence-based educational research in China

Notwithstanding the growing awareness of conducting evidence-based educational research among many Chinese scholars, some remain skeptical about its feasibility of operation and effectiveness in informing practices. Some aspects of skepticism can be summarized as follows.

The definition of "evidence-based". Some scholars argue that "evidence-based" could be understood from both narrow and broad perspectives. For example, Shi (2017) argues that if "evidence-based" is understood from a broad sense, namely researchers collecting and observing relevant data to propose and validate hypotheses, then Chinese academia has all along been engaged in this aspect of "evidence-based." However, Shi states, if "evidence-based" is interpreted from a narrow perspective, namely using procedural, operational, and quantitative measures to draw conclusions for certain educational and social phenomena, that is not what most Chinese educational researchers have been doing.

At one with such a possibly broad definition of "evidence-based," quite a number of Chinese scholars have published articles that call for putting evidence-based research in perspective. For instance, Li (2018) warns that evidence-based research should not be regarded as "statisticalism." He notes that given the growing emphasis on conducting evidence-based research, there is a tendency to equate "evidence-based" with numbers and statistics, which he terms as "statisticalism." He argues that such a mentality and approach "simplifies the research process into the operational process of statistics, and adopting an exclusionary attitude towards educational ideas, values and emotions that are difficult to quantify." He summarizes the features of statisticalism as "focusing on the method while neglecting the problem, focusing on data while neglecting the facts, focusing on the format while neglecting the thought." As such, he argues that the connotations of empirical research and evidence-based research should be reexamined.

Such an idea is echoed by Wang (2019) who argues that evidence-based research tends to fall short in discussing value-related issues in various fields, including education. He argues that value

is concerned about the future of mankind, specifically the choices that people ought to make in face of interest conflicts, while evidence-based research tends to focus on the present reality that overlooks the value of subjectivity and imagination of human beings. As such, he holds that speculative educational research should serve as a prerequisite for evidence-based educational research, and also be inclusive of findings generated from evidence-based research.

In the same vein, Pan (2018) holds that it is necessary to distinguish similar concepts around "evidence-based research." He argues that evidence-based research should not be equated with quantitative research, while he admits that quantity underpins a large proportion of evidence-based research. He believes that rigorous qualitative approaches that validate certain hypotheses also lead researchers to real evidence and contribute to resolving educational problems. He especially emphasizes the value of ethnographic research that generates new theories. Therefore, he holds that due to the complex and multidimensional nature of educational problems, educational research should not be reduced or narrowed into empirical research or quantitative studies.

Approaches to collecting real evidence. Some other scholars have cast doubt on the feasibility of collecting "real" evidence in educational settings (Liu, 2017; Yang et al., 2019). For instance, Yang et al. (2019) maintain that while the formal procedures of conducting evidence-based research such as experiments or quasi-experiments are understood by the research community, realistic educational settings may not always allow such procedures to be rigorously fulfilled given the relevant conditional constraints. They enumerate several frequently encountered difficulties during the process of data collection, such as limited access to high-quality data, the costs of collecting longitudinal data, the unclear procedures and standards for collecting data, and the low quality of data collected due to insufficient training of researchers. The three authors especially highlight these plights for individual researchers engaged in data collection and point out their negative effects on the analysis of data collected and conclusions drawn.

Similarly, Liu (2017) argues that because evidence-based research generally holds that any concept is meaningful only when it can be reduced to empirical contents that individuals experience, the reality is that education in and of itself involves plenty of complexity and even abstract ideas that are not easily reducible. Therefore, he casts doubts on the credibility and validity of evidence-based approaches (e.g., using scales) to effectively capture the essentials of certain educational connotations (i.e., morality and happiness). Many important concepts in educational science such as educational purpose, ethics, and values, according to Liu, are difficult to quantify and observe. Therefore, Liu also believes in the importance of experience in underpinning the establishment of a set of educational research paradigms.

Tian (2018) holds a similar skeptical attitude toward the quality of instruments used to collect data among Chinese scholars, especially among students. The weakness in mastery of

methodology, the lack of sufficient conditioning support, the unsound academic climate, and misbeliefs about research paradigms are the top four factors he identifies that undermine the high-quality implementation of evidence-based research in China. Furthermore, he argues that those four factors have complex interactions with other variables such as gender, university, academic years, and program types, thus displaying a landscape of complexity and complication.

Evidence-based reform in the U.S. and UK

The situation of educational research in the U.S. and UK has many similarities to that in China. Educational research has long used a wide variety of methods to learn about educational policy, practice, and outcomes. These include quantitative research, such as experiments, correlational studies, and surveys, but also qualitative designs, such as systematic observation, surveys, and interviews, as well as mixed-method investigations. All of these types of research continue in full force today throughout the West, as they do in China.

However, what has substantially changed over the past 20 years is a strong policy focus in government on using the tools of educational research to evaluate practical programs for schools. This type of research does not only seek wisdom, it is also intended to produce evidence of effectiveness to directly inform educators and policymakers about the likely impacts on student outcomes of using one program versus another, or one policy versus another. This type of research asks questions such as, "What is the impact of a specific technology program, compared to standard practice, on student achievement in biology? What is the effect of a particular strategy for grouping students on literacy achievement? What is the effect on student learning of teaching teachers to use particular forms of cooperative learning in their mathematics classes?"

For these kinds of questions, quantitative methods are required, as these are all "how much" questions. Ideally, schools, teachers, or students are assigned at random to use an experimental program, or to remain in a traditionally taught control group. If random assignment is not possible, schools, teachers, or students may be matched at the beginning of a "quasi-experiment." Such studies usually use mixed-methods, including observations or ethnography, to add understanding to the study, but the main emphasis is on the quantitative comparison of one group to another, in learning or other outcomes.

The reason for policies favorable to experiments is political support for research that provides practical solutions to important problems of education. Policymakers are more eager to improve student outcomes than to fill educational journals. Educational research is still underfunded in all countries, but the progress of studies capable of providing reliable evidence of effectiveness of particular, replicable methods has brought historic levels of funding into educational research. As noted earlier, the approach taken by this type of experimentation derives substantially from

medicine, which has made extraordinary progress due to an insistence on rigorous evaluations of new medicines and procedures.

The following sections of this article focus on new policies and practices regarding educational research and dissemination of proven programs in the U.S. and UK, and implications of these developments for research, practice, and policy in China.

Every Student Succeeds Act evidence standards

In recent years, a major policy initiative in the U.S. has created new possibilities for evidence-based reform in education. The Every Student Succeeds Act (ESSA) became the main federal education law in December, 2015. It includes definitions of strong, moderate, and promising levels of evidence supporting education programs. The definitions are as follows:

> (i) demonstrates a statistically significant effect on improving student outcomes or other relevant outcomes based on—
>
> (I) *strong evidence* from at least one well-designed and well-implemented experimental study;
>
> (II) *moderate evidence* from at least one well-designed and well-implemented quasi-experimental study; or
>
> (III) *promising evidence* from at least one well-designed and well-implemented correlational study with statistical controls for selection bias.

The evidence standards are not self-enforcing but must be broadly supported if they are to have any impact. The purpose of this article is to discuss what China could do to progressively increase the role of evidence in its own education policies and practices, with the goal of increasing school effectiveness and student outcomes on a substantial scale.

Essential steps in evidence-based reform in education

Finding out what works best in every subject and grade level

Perhaps the most important requirement for evidence-based reform is that there be a substantial set of replicable programs and practices with clear evidence of effectiveness. For example, imagine that a government wanted to improve the reading achievement of young children. Experts on reading might be asked to design and initially evaluate various methods for teaching reading to students who are 5 or 6. The most promising methods might then be evaluated using a randomized experiment. For example, 50 schools might be identified to participate in an experiment; 25 of the schools would be assigned at random (e.g., by coin flips) to either use the new program or to continue with usual practices. Students would take pretests before the new program begins and then would take posttests at the end of the school year. Researchers would collect the pretests and posttests and analyze them statistically to find out which group learned more. If the students who

used the new program scored significantly higher than the control group, who used usual practices, then the new program would be known to be effective, especially if similar experiments are repeated elsewhere and are also found to be successful.

This design, comparing schools using a new program to those using usual methods, could be applied in any subject or grade level. Whether the subject is high school science, middle school mathematics, elementary English, or any other subject, the same basic evaluation methods would work. Over a few years, government leaders could see which programs work better than usual programs, and they could promote the use of these successful methods.

Policies in the U.S. and UK have invested in the development and evaluation of many programs. These government initiatives are as follows.

Investing in Innovation and Education and Innovation Research. A major advance in the creation of a large set of proven programs was the U.S. Investing in Innovation (i3) grant program, established in 2009 to fund the development, evaluation, and scale-up of proven programs at all levels of education, pre-K to 12, using a "tiered funding" strategy. Programs that already had strong evidence of effectiveness could qualify for scale-up grants, those with some evidence could qualify for smaller validation grants, and those with a good but untested idea could receive smaller development grants. By its final year in 2016, i3 had funded 171 projects to do randomized evaluations of a wide variety of programs. i3 has now been replaced by the Education and Innovation Research (EIR) grant program which also utilizes a tiered evidence structure.

Institute for Education Sciences. Institute for Education Sciences (IES) has been supporting development and rigorous evaluation of scalable programs in the U.S. since 2002 and, in the process, has greatly built up the national pool of capable evaluators on which i3 and EIR rely. IES also funds research and development in a tiered structure.

Education Endowment Foundation (England). A major effort patterned on i3 is taking place in England, where the current government allocated substantial funding to a private foundation to commission evaluations of promising programs for primary and secondary schools. This Education Endowment Foundation has funded many third-party, mostly randomized evaluations of a wide variety of programs.

Making educators and policymakers aware of proven programs and practices

It is not enough to know which educational programs work in each subject and grade level. This information must be shared with educators, education leaders, and policymakers, in a way that is readily accessible to all. This requires that trusted government or nongovernment (e.g., university)

experts interpret technical reports on evaluations of educational research, describe programs and their evidence base clearly, and keep this information up to date.

What Works Clearinghouse (www.whatworks.org). Education policies cannot support evidence-based practice unless there is some agreement on which specific programs have clear evidence of effectiveness. In the U.S., a flagship initiative of IES is the What Works Clearinghouse (WWC), established to review research on practical programs in many areas. The WWC now has reports on research in reading, math, science, high school graduation, and other areas, and it produces readable "practice guides" on several topics.

Best-Evidence Encyclopedia (www.bestevidence.org). The Best-Evidence Encyclopedia (BEE) is a free website created and routinely updated by our group at Johns Hopkins University. It primarily contains full-scale academic reviews of research on the effectiveness of programs in specific areas. For example, it has reviews on elementary and secondary reading, reading for struggling readers, elementary and secondary math, elementary and secondary science, and prekindergarten programs.

Evidence for ESSA (www.evidenceforessa.org). Evidence for ESSA is another free website from our group at Johns Hopkins, but unlike the BEE, it is targeted primarily at educators, not academics, and is aligned with the ESSA evidence standards. For this reason, our group that created the BEE developed a new website solely committed to communicating to educational leaders information on individual programs that do or do not meet the ESSA definitions for strong, moderate, or promising evidence of effectiveness.

Evidence for ESSA has completed reviews of programs for reading and math, grades pre-K to 12, as well as social-emotional learning and attendance, grades K-12. About 100 reading and mathematics programs meet the "strong," "moderate," or "promising" standards, with the largest number in the "strong" category. It has located 24 social-emotional learning programs and 8 attendance programs meeting ESSA standards. As of January 2020, Evidence for ESSA had more than 100,000 unique users, adding another 800 each week.

BEiB and blogs. Begun in a collaboration between Johns Hopkins University and the University of York in England, BEiB (https://beibindex.wordpress.com/) is sent at no cost every 2 weeks to tens of thousands of subscribers in the U.S. and UK. It contains brief articles on recent studies, policies, and other news of interest to researchers, educators, and policymakers interested in evidence-based reform. More recently, NNU has created a Chinese version of BEiB focused on the Chinese mainland (https://blog.educaixa.com/es/evidencias-educaixa1) and the Chinese University of Hong Kong created a version focused on overseas Chinese (https://cuspbeb.com/en/). In both cases, these versions focus primarily on research carried out in China or other Chinese-speaking regions, although they do translate particularly applicable articles from the U.S./UK version with

adaptations to the Chinese context. The U.S./UK version of BEiB similarly publishes articles adapted from the Chinese versions.

In addition to BEiB, Robert Slavin writes weekly blogs (https://robertslavinsblog.wordpress.com/) on a wide variety of topics relating to evidence-based reform, methodology, and policy. Both the blogs and the U.S./UK version of BEiB include an index, to enable readers to easily find articles from previous editions on all topics.

Providing resources and incentives to enable school leaders to implement proven programs and practices

Knowing which programs work best is rarely enough to cause educational leaders to embrace them. Educational innovations usually have initial costs (for materials, hardware, software, and professional development) and educators want assurance that adopting innovations is encouraged by their leaders. For these reasons, government must make a commitment to provide resources and encouragement to school leaders willing to try out new programs.

In the U.S., the ESSA law provides both financial support and encouragement for schools to use to implement proven programs. This provision is primarily directed at very low-achieving schools, but the conversation about evidence has led several states to insist that schools use evidence-based programs if they seek certain state funds, not just federal funds. This is just beginning to happen, so the larger impact of encouragement and assistance to select proven programs is not known. However, previous incentive programs, such as the National Diffusion Network in the 1980s, Comprehensive School Reform in the 1990s and i3 from 2010 to 2018 have shown that government encouragement and modest funding can cause thousands of schools to adopt proven or promising programs.

In the UK in 2018, government almost implemented a program that would have provided significant funding to help schools adopt proven programs. At the last minute, it was canceled, but this idea may reappear in the future.

Creating policies and systems to continuously add to knowledge and effective practice

Implementing evidence-based practices on a large scale could certainly be accomplished. China lends itself to such a strategy because it is such a large country with most citizens speaking the same language. If Chinese researchers create and successfully evaluate a new approach to writing or biology or geometry, this could potentially be used in hundreds of thousands of schools. However, national and local policies would need to support the entire process of development, piloting, evaluation, and dissemination.

Some progress could be made right away by evaluating existing programs if the evaluations show promise, disseminating proven approaches. However, to rapidly build capacity to do this

work, the Chinese government would need to build expertise at all levels of the innovation–evaluation–dissemination process. For example, there would need to be many highly skilled people able to create innovations or adapt them from abroad. There would need to be many more skilled evaluators to evaluate new programs. There would need to be many more people capable of continually summarizing evidence on effective programs and making it easily available and easy to interpret for educators. There would need to be many people able to use state-of-the-art, culturally appropriate methods to train teachers in implementing proven programs.

The need for new ways to solve fundamental problems in education is never-ending. Even if very effective approaches, proven in research, are already established, there are always new discoveries that can further improve outcomes. Besides, conditions change over time, and research and development need to be able to respond to this change. Also, different instructional approaches may be needed for different types of schools, such as urban and rural schools, or schools mostly serving children of well-educated parents and those serving many children of less-educated parents. Research in education should not be seen as a search for a single best method for all schools, but as a continuing process of innovation and evaluation to progressively add to knowledge about effective practices and policies.

Further, the purpose of education research should not just be to produce higher test scores or other quantitative outcomes. Innovation and evaluation need to incorporate study of instructional processes, student and teacher attitudes, and pragmatic difficulties and solutions. Over time, studies of improving teaching strategies, materials, and technology should help researchers and educators alike better understand how learning takes place and how change can be supported and maintained.

If the Chinese government decides to improve school effectiveness with proven programs, it might identify a few regions in which to test out policies and practices, to learn how to make these ideas work in the Chinese context. There is every reason to believe that evidence-based reform could be successful in China, but there is no reason to believe that the path to this goal will be the same in China as it has been in the U.S., the UK, and other countries.

China has already created a successful educational system, with outstanding results, as shown on international tests such as PISA. However, China is faced with unique educational issues, such as how to better incorporate quality and competence-oriented elements in the current educational system that prioritizes standardized tests, and how much workloads should be reduced for students in order to keep a better balance between learning and whole-person development. Evidence especially matters in informing the government and schools as to which reform measures are more effective in achieving overall educational reform goals. As education becomes increasingly important in building economic security, every country needs to consider how to build education systems ready for the future. Evidence has been the key to world progress in medicine, agriculture, technology, and other fields. Why should it not be the key to improving education?

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Contributorship

Robert Slavin was responsible for writing the Abstract, the bulk of the main body, finalizing the paper and responding to reviewers' comments. He covered various aspects of the paper including theorizing and analyzing how evidence-based education could advance Chinese education, why evidence should matter, evidence-based reform in U.S. and UK, etc. Alan C. K. Cheung contributed by identifying, analyzing, and responding thoroughly to skepticism of evidence-based educational research in China based on relevant Chinese and English literature. Tengteng Zhuang teased out literature and materials on evidence-based reform in Chinese education in recent years, including searching for and summarizing data in Chinese literature and writing relevant sections of the paper.

References

Bridgeland, J., & Orszag, P. (2013). *Can government play moneyball*? http://www.theatlantic.com/magazine/archive/2013/07/can-government-play-moneyball/309389/

Bucks, S., & McGee, J. (2015). *Why government needs more randomized control trials: Refuting the myths*. Laura and John Arnold Foundation.

Gueron, J. M., & Rolston, H. (2013). *Fighting for reliable evidence*. Russell SAGE.

Haskins, R. (2014, December 31). Social programs that work. *The New York Times*.

Hemsley-Brown, J., & Sharp, C. (2003). The use of research to improve professional practice: A systematic review of literature. *Oxford Review of Education*, *29*(4), 449–470. https://www.sohu.com/a/147645112_498166

Kolada, G. (2013, September 2). Guesses and hype give way to data in study of education. *The New York Times*. http://www.nytimes.com/2013/09/03/science/applying-new-rigor-in-studying-education.html?_rD0

Li, J. (2018). Education empirical research should not fall into the mode of "statisticalism" [in Chinese]. *Journal Higher Education*, *39*(11), 64–70.

Liu, X. (2017). How to carry out evidence-based research and what confuses evidence-based educational researchers—Implications from the 2nd ECNU national forum on evidence-based educational research [in Chinese]. *Modern Distance Education Research*, *147*(3), 18–25.

Pan, H. (2018). Limitations and transcendence of educational empirical research [in Chinese]. *Contemporary Education Review*, *8*, 10–13.

Shi, J. (2017). How should we perceive and implement evidence-based educational research? [in Chinese]. https://www.sohu.com/a/147645112_498166

Slavin, R. E. (2008). Evidence-based reform in education: What will it take? *European Educational Research Journal*, *7*(1), 124–128.

Slavin, R. E. (2013). Overcoming the four barriers to evidence-based education. *Education Week*, *32*(29), 24.

Slavin, R. E. (2017). Evidence-based reform in education. *Journal of Education for Students Placed at Risk (JESPAR)*, *22*(3), 178–184.

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

Slavin, R., & Cheung, A. C. K. (2017). Lessons learned from large-scale randomized experiments. *Journal of Education for Students Placed at Risk (JESPAR)*, *22*(4), 253–259. https://doi.org/10.1080/10824669.2017.1360774

Tan, C. (2017). A review of empirical research on ideological and political education in China in the past 10 years—Content analysis of 1399 papers in five journals [in Chinese]. *Studies in Ideological Education*, (9), 18–23.

Tian, X. (2018). What restricts the development of positive research in education?—A positive study based on Chinese graduate students [in Chinese]. *Education Science*, *34*(2), 53–59.

Wang, W. (2019). Speculative educational research and empirical educational research: From separation to symbiosis [in Chinese]. *Educational Research*, *476*(9), 139–148.

Wang, X. (2017). On empirical research of higher education—An analysis based on periodicals [in Chinese]. *Heilongjiang Researches on Higher Education*, *284*(12), 5–9.

Weiss, C., Murphy-Graham, E., Petrosino, A., & Gandhi, A. (2008). The Fairy Godmother—And her warts: Making the dream of evidence-based policy come true. *American Journal of Evaluation*, *29*(1), 29–47.

Yang, Z., Han, R., & Lin, X. (2019). The basic norms, realistic dilemmas and improvements of data collection in educational empirical research [in Chinese]. *Educational Measurement and Evaluation*, (8), 14–20.

Yuan, Z. (2017). Chinese education needs evidence-based research [in Chinese]. *Journal of the Chinese Society of Education*, (2), 1–2.