

**INTERNATIONAL PERSPECTIVES IN  
EDUCATIONAL EFFECTIVENESS RESEARCH**

**Edited by James Hall<sup>1</sup>, Ariel Lindorff<sup>2</sup> and Pamela Sammons<sup>2</sup>**

**<sup>1</sup> Southampton Education School, University of Southampton,  
UK**

**<sup>2</sup> Department of Education, University of Oxford, UK**

## Table of Contents

<b>1 Introduction</b> .....	1
James Hall, Ariel Lindorff and Pamela Sammons	
<b>Section 1: Setting the scene for international perspectives in Educational Effectiveness Research: Historical context, theory, and extensions</b>	
<b>2 International perspectives in educational effectiveness research: A historical overview</b> .....	10
Ariel Lindorff, Pamela Sammons and James Hall	
<b>3 Developing and testing theories of educational effectiveness addressing the dynamic nature of education</b> .....	40
Leonidas Kyriakides, Bert Creemers and Anastasia Panayiotou	
<b>4 The Fifth Phase of Educational Effectiveness Research: The philosophy and measurement of equity</b> .....	87
Anthony Kelly	
<b>5 Extending educational effectiveness: The middle tier and network effectiveness</b> .....	125
Daniel Muijs	
<b>6 Extending Educational Effectiveness: A critical review of research approaches in International Effectiveness Research, and proposals to improve them</b> .....	150
David Reynolds, Anthony Kelly, Alma Harris, Michelle Jones, Donnie Adams, Zhenzhen Miao and Christian Bokhove	
<b>7 Policies and practices of assessment: A showcase for the use (and misuse) of International Large Scale Assessments in Educational Effectiveness Research</b> .....	180
Eckhard Klieme	
<b>Section 2: Examples of Educational Effective Research from around the globe</b>	
<b>8 Educational Effectiveness Research in Africa: The case of the Democratic Republic of the Congo (DRC)</b> .....	226
Gratien Mokonzi Bambanota, Jan Van Damme, Bieke De Fraine, Paul Vitamara Masimango, Gaston Kimbuani Mabela, Augustin Mukiekie Tshite, Stanislas Maroyi Lukula, Oscar Gboisso Asobee and Jean Paul Bela Legono	

<b>9 School and teacher value added performance and the relationship with teacher professional development in mainland China.....</b>	<b>256</b>
Sally M. Thomas	
<b>10 Three decades of Educational Effectiveness Research in Belgium and the Netherlands: Key studies, main research topics and findings.....</b>	<b>282</b>
Marie-Christine Opdenakker	
<b>11 The impact of socioeconomic segregation in U.S. high schools on achievement, behavior, and attainment and the mediating effects of peers and school practices.....</b>	<b>357</b>
Gregory J. Palardy	
<b>12 Leadership for Learning in diverse settings: School leaders setting the agenda in Australia and New Zealand.....</b>	<b>406</b>
Tony Townsend, Mere Berryman, David Gurr and Lawrie Drysdale	
<b>13 A national evaluation of kindergarten outcomes: Findings from Uruguay.....</b>	<b>448</b>
Alma Y. Lopez and J. Douglas Willms	
<b>14 Continuing towards international perspectives in Educational Effectiveness Research.....</b>	<b>474</b>
James Hall, Pamela Sammons and Ariel Lindorff	

## 2 International perspectives in educational effectiveness research: A historical overview

Ariel Lindorff, Pamela Sammons and James Hall

**Abstract** This chapter provides a concise narrative overview of the development of international perspectives in Educational Effectiveness Research (EER), contextualised within shifts towards globalisation in education policy including the increased prominence of international large-scale educational assessments (ILSA) in shaping international policy discourse. We take into consideration the broad range of ways in which international perspectives have manifested within the EER field, ranging from published international comparative studies, diverse single-country studies, reviews and meta-analyses, to the “softer” but nonetheless influential intellectual interchange facilitated by the formation of an international community of researchers, policymakers and practitioners.

We begin by reflecting on the earliest research in the field, which took place in a limited number of countries, but – we suggest – facilitated the emergence of a nascent international dialogue. The subsequent formation of international organisations – particularly the International Congress for School Effectiveness and Improvement (ICSEI) – bringing together EER researchers, practitioners and policymakers, has given this dialogue a more concrete forum over the last 30 years. We document how the search for the “big picture” of educational effectiveness on an international scale has informed a more sophisticated and nuanced set of international perspectives in EER, via increasingly diverse single-country studies, reviews, meta-analyses, international comparative studies, and the development and empirical testing of EER theories. We note recent trends towards a combination of increased internationalisation, increased synergy with and reciprocal influence upon ILSA, and increased methodological diversity in EER. Finally, we offer some concluding reflections and recommendations for future investigations and further development of the field.

## 2.1 Introduction

The origins of the modern field of Educational Effectiveness Research (EER) lie in debates surrounding the question, “Do schools make a difference?”. One result was the development of the field of School Effectiveness Research (SER) – a field whose research is today encompassed within the more-broadly focussed EER. Extensive discussion of the evolution from SER to EER have been provided by Teddlie and Reynolds (2000), Creemers, Kyriakides & Sammons (2010), Chapman, Reynolds, Muijs, Sammons, Stringfield and Teddlie (2016), and Sammons, Davis, & Gray, (2016). Over nearly 50 years, EER has evolved to become an intellectually coherent and widely recognised international field that has moved beyond a focus just on schools and their effects to now ask and answer increasingly complex and nuanced research questions about effects of teachers, classrooms, institutions, networks, and systems. It can be thought of in the present day as the broad endeavour to address questions including:

- How, why, when, where and for whom do education systems, teachers and schools make a difference with regard to a variety of student outcomes?
- What are the most appropriate designs and methodologies to undertake such research?
- How can findings from EER provide evidence and information that can be used to improve learning and life chances for all students, particularly disadvantaged and ‘at risk’ groups in different cultural contexts?

Rather than being concerned merely with “what works” (a notion more commonly associated with Randomised Controlled Trials, RCTs), EER has historically adopted an ecological perspective that stresses the importance of investigating “natural” variation at different levels of education systems, neighbourhoods and schools (or other institutions), to shed light on issues of equity, quality, and susceptibility to intervention via improvement efforts. Although the roots of EER can be traced back to work conducted in a small number of countries (e.g. the USA, the UK, the Netherlands, and Australia), the field has become increasingly internationally diverse over time. This change has also been required given that questions of educational quality and equity are of universal relevance (even if the answers to these questions and the specific circumstances surrounding them may

vary by culture and by country), and because education is increasingly seen as an important policy lever to achieve social and economic aims.

Although this chapter (and the volume within which it is situated) focus on EER, both acknowledge the contribution of educational improvement research towards the development of EER in three areas simultaneously: its internationalisation, its methodologies, and its theoretical perspectives (e.g. Harris & Chrispeels, 2006). Although educational improvement and educational effectiveness developed from different original perspectives and approaches, both share an emphasis on what education systems, institutions and teachers can do to foster positive outcomes for their students, and the two fields have grown together over time and been reciprocally informative. For example, Chapter 6 (Reynolds et al., n.d.) uses the acronym EEIR to capture this shared focus on both effectiveness and improvement and both these terms feature within the name and objectives of the ICSEI organisation. The benefits of this are two-fold and mutual. On one hand, the more we learn about the characteristics and processes associated with positive student outcomes, the more insight we have about what can be improved and how (Mortimore, 1998). On the other hand, the more we understand the contexts and cultures within which education operates locally and globally, the better we can adapt our approaches to gain richer and fuller insight regarding the quality and equity of education around the world.

The internationalisation of EER also needs to be contextualised against the backdrop of the increased globalisation of education policy and practice in general. The following sections manifest this reality with a historical account of the international perspective in EER alongside broader contemporary trends towards globalisation in the policy sphere as well as the increasing prevalence and influence of international large-scale assessments (ILSA) and surveys. With the increasing attention of policymakers on the outcomes of international assessments, and their country's league table positions in such assessments and surveys, it is hardly surprising that education policy itself has become increasingly globalised. That is, policymakers have been increasingly concerned with the ability of schools to prepare children and young people to participate in the "global economy", and policy transfer between countries has become increasingly common (for

examples see Chapters 6 and 10 in this volume by Reynolds et al. and Oppdenakker, n.d., respectively). Many supranational organisations such as the OECD and World Bank continue to accord educational outcomes as key benchmark indicators and set ambitious global educational goals concerning access, participation, and increasingly a focus on quality, equity and improvement of educational outcomes that increasingly shape policy debate and influence national agendas.

A contemporary international example of the increasingly globalisation of educational policy and practice includes the current attempts to transplant, adapt and apply various East Asian approaches to teaching and learning into other settings. In the USA, this has included a push to adopt Singaporean mathematics teaching approaches and textbooks, with a fairly limited evidence base to support the effectiveness of this (e.g. Jaciw et al., 2016). In England, a similar interest in East Asian approaches has emerged with a large amount of resources spent in 2016 to encourage schools to use textbooks and teaching approaches modelled on those used in Shanghai and Singapore. The evidence base for such policy transfer, however, was not strongly rooted in robust empirical evidence; such evidence has been relatively limited and recent (e.g. Hall, Lindorff, & Sammons, 2016; Jerrim & Vignoles, 2015; Lindorff, Hall, & Sammons, 2019). That is, the perceived effectiveness of an educational approach in one cultural setting (East Asian) has often been taken as convincing evidence of its potential to shape policy initiatives intended to change educational practices in another context (Western systems). Raising mathematics has been identified as of particular policy interest in some countries with lower performance in international assessments. However, this assumption of simple cross-country transference is highly problematic, given the complexity of the diverse social, cultural and political contexts within which education in a particular setting is situated.

An additional development in the education policy sphere has been an increasing emphasis on evidence seen as “scientific”<sup>2</sup>, particularly

---

<sup>2</sup> Note: From the authors’ perspective, scientific evidence can take multiple forms. In presenting “evidence seen as ‘scientific’” in a policy context, we do not propose that experimental evidence is more or less valuable, scientific or valid than observational evidence, merely that the former has – in recent years – appeared to gain prominence and perceived credibility over the latter in non-academic spheres.

priority given to experimental studies which are often lauded as being a ‘gold standard’ for producing policy-relevant findings. This is reflected in a shift in funding priorities within countries where this emphasis is especially dominant, and the proliferation of Randomised Controlled Trials (RCTs) and/or Quasi-Experiment Designs (QEDs) to shape policies and practice. For example, both the Education Endowment Foundation (EEF) in England and the What Works Clearinghouse (WWC) in the USA have been influential drivers of evidence-based practice. Both place explicit priority on experimental designs as the most ‘trustworthy’ form of evidence (Education Endowment Foundation, 2018; What Works Clearinghouse, 2017). Perhaps as a consequence, there has been some complication and confounding of terminology that has long been used in EER, and some regional trends of decline in its prevalence. For example, the Society for Research on Educational Effectiveness (SREE), founded in 2005, states that it “brings together individuals interested in the use of causal inference to improve educational practice” (SREE, 2019). Without devaluing experimental (or quasi-experimental) research or the evidence arising from it, it is important to consider the implications for the field of EER of having, in effect, two uses of the word “effectiveness” in quite distinct ways – one grounded in experimental manipulation, the other grounded in observation and analysis of natural variation in nested educational systems at different levels (e.g. national, regional, neighbourhood, school, classes and student). Both may contribute valuable evidence within and beyond the scope of EER, and indeed the more inclusive EEIR that encompasses improvement initiatives as well as studies of educational effectiveness, but the ambiguity of the term “effectiveness” may risk confusion of knowledge bases and professional and academic communities of practice.

Given the growing globalisation of education policy and practice, EER has undergone a dramatic and rapid transformation over the last few decades as comparison of major overviews illustrate (Teddlie & Reynolds, 2000; Chapman et al., 2016), and a greater emphasis on methodological issues and advancement of the field (Creemers, Kyriakides & Sammons, 2010). A considerable part of this transformation has involved looking outward from the few places from which the earliest EER (then within SER) emerged to other country contexts and the use of more sophisticated research designs and comparative

perspectives. In the sections that follow, we endeavour to trace the history of international perspectives in EER, taking into account the wide variety of studies and developments that together comprise the broad range of manifestations and interpretations of an “international perspective”. Rather than limiting our account to comparative international studies, we consider also the international conversations and collections sparked by single-country studies, meta-analyses and syntheses across diverse contexts, as well as steps toward internationalisation of EER that have not been limited to formal academic literature but have arisen from the development of an international community of researchers, policymakers and practitioners.

## **2.2 Early EER: The beginning of an international dialogue**

The earliest research on the effectiveness of schools emerged as a response to studies in the USA by Coleman et al. (1966) and Jencks et al. (1972), both of whom found that the “effects” of schools on student outcomes were relatively small in comparison to the effects of student background characteristics; a conclusion also reached by the then contemporary Plowden Report in the UK (Plowden & Central Advisory Council for England, 1967). The implications drawn, then, were that schools did not make much of a difference in terms of student achievement, and schools could therefore do little to mitigate the effects of disadvantage. In reaction to this came studies in the USA (Brookover et al., 1978; Edmonds, 1979; Weber, 1971) and UK (Rutter, Maughan, Mortimore, & Ouston, 1979). These not only indicated that schools had an effect but provided evidence regarding what factors seemed to account for variation between schools in student outcomes. Reynolds (2000) suggested that school/educational effectiveness studies in this early phase showed strongly “ethnocentric tendencies” (p. 232): studies were generally done in a small selection of settings and reviews of the relevant literature tended to emphasize heavily – if not exclusively – the research conducted in the author’s own country.

Compared to studies examining the effectiveness of schools, Teacher Effectiveness Research (TER) arguably began much earlier. The history of this branch of EER can be traced back to studies of

teachers perceived to be effective in the late 19<sup>th</sup> century (e.g. Kratz, 1896) and the beginnings of systematic observation (e.g. Stevens, 1912), but it was not until the 1960s that investigations of the relationships between teacher behaviours and student learning outcomes began to emerge (e.g. Gage, 1963). Teacher effectiveness has its roots in North America and shares some of the same general approaches as school effectiveness research in its search for effectiveness factors, albeit while focusing on the classroom rather than the school as a whole (Chapman et al., 2016; Creemers, 1994). As with SER though, TER studies too were generally done in a single context and tended to prioritise, review and acknowledge literature from within that context (e.g. Doyle, 1977).

The focus on within country and culture context in the early days of both school and teacher effectiveness research was not, however, to the exclusion of any international perspective. Instead, these early studies did initiate some degree of dialogue spanning beyond national boundaries. The Coleman (1966) and Jencks (1972) studies elicited responses that were not limited to their country of origin, and the seminal studies by authors including Edmonds (1979), Rutter et al (1979), and (Mortimore et al. (1988) were, and continue to be, frequently cited in literature from outside of their original settings. In that sense, although early reviews of educational effectiveness literature did not always clearly identify or discuss context or national differences, international perspectives were already taking shape – albeit not always explicitly and to a somewhat limited extent – in the form of that international intellectual exchange in the academic sphere. As Mortimore (2001) put it, “...The early existence of independent research projects in two countries asking similar questions and drawing, to a certain extent, on similar methodologies demonstrated the potential for further global investigations” (p. 236).

### **2.3 The contribution of the International Congress for School Effectiveness and School Improvement: An international community of research and practice**

An international community was already starting to take shape only about a decade after the earliest school effectiveness studies. 1988 marked a pivotal moment in the history of EER: The first meeting of the International Congress for School Effectiveness took place in London, with 14 participating countries (see Townsend, 2007, for a more comprehensive narrative account of the origins of this organisation). The Congress was intentionally convened not only as a community of researchers, but as an opportunity to bring researchers, policy-makers and practitioners together to share ideas and to shape and improve research and practice. Not long after that first meeting, the name of the organisation was updated to the International Congress for School Effectiveness and Improvement (ICSEI) to reflect the ambition to promote convergence of the school effectiveness and school improvement fields. Soon afterward, the journal *School Effectiveness and School Improvement* was established (and associated with ICSEI), with its first issue published in 1990. It has since provided one of the key outlets for research in the field of EER. As of 2019, and as a marker of the extent to which the international EER community has diversified and expanded, ICSEI members come from over 50 countries and a deliberate effort is made by ICSEI to ensure that annual meetings of the organisation take place on different continents each year to promote an inclusive culture and to facilitate the involvement of practitioners and policymakers who might be less able to attend international or distant meetings.

ICSEI is by no means the only forum for the EER and school improvement community, though it is an influential one. Over time, other groups have formed, including the Society for Research in Educational Effectiveness (SREE) with its affiliated journal, the *Journal for Research in Educational Effectiveness* (first published in 2008), the Educational Effectiveness special interest group (SIG 18) in the European Association for Learning and Instruction (EARLI), the Educational Effectiveness and Quality Assurance network of the European Conference on Educational Research (ECER), the School Effectiveness and School Improvement SIG in the American Educational

Research Association (AERA) and the Educational Effectiveness and Improvement SIG in the British Education Research Association (BERA). ICSEI, however, is notable in that it has taken as part of its core aims facilitating research and improvement activities through international collaboration and knowledge-sharing, forging links between educational effectiveness and educational improvement, and bridging between research, practice and policy in different international contexts across the globe (e.g. see Schildkamp, 2019, for text from the 2019 President of ICSEI that reflects these aims).

ICSEI has also sought to support international advancement of EER methodology through the creation of the MoREI (Methods of Researching Educational Effectiveness and Improvement) network. This has supported a number of cross-country international collaborations and an interest in developing measures that can be used in different contexts to facilitate comparative research. The International Schedule for Teacher Observation and Feedback (ISTOF; Teddlie et al., 2006; Muijs et al., 2018) provides an example of a research collaboration that developed a new instrument based on EER and TER research evidence and expert opinion from 20 participating countries. It is intended to support both research and improvement activities reflecting ICSEI's core aims. We consider the ISTOF project in more detail in Section 6.

Country reports presented at some of the ICSEI meetings have highlighted the variety of EER studies taking place around the world, and have been published as collections with bridging commentary (e.g. Creemers, Peters, & Reynolds, 1989; Reynolds, 1996; Townsend, Clarke, & Ainscow, 1999). These collections and the editors' reflections on them contributed, at the time of their publication, a new form of international perspective, looking across contexts and cultures for thematic contributions to as well as similarities and differences in research, policy and practice and the interchange between them. The countries and regions of the world represented in these collections of country reports expanded; by the time Townsend, Clarke and Ainscow's *Third Millennium Schools* (1999) was published, 20 countries were represented, covering all major regions of the globe.

Efforts to create a more inclusive and international EER community have been concerted and ambitious, but their success has

arguably been tempered by practical challenges. ICSEI, for example, has expanded its membership but remains to some extent dominated by researchers and policymakers from North America, the UK and Continental Europe. As noted above, the organisation intentionally avoids holding its annual meetings exclusively in these regions, ensuring that it rotates between different regions and includes both high- and low-income host countries, but even with subsidised participation for participants from low-income countries and explicit policy governing the variation in host countries, there are doubtlessly still barriers to participation for those coming from under-funded institutions and low-income nations.

#### **2.4 The contribution of single-setting studies: A growing body of evidence**

The ICSEI country reports discussed in Section 3 constitute just a sample of the single-country EER studies that have been, and are being, conducted, but they also provide an indication of the spread of EER internationally. By the end of the 1900s although EER studies remained concentrated within the countries in which the field was first established, the settings in which single-country EER studies were taking place had diversified considerably. The subsequent contribution of these studies to an international perspective in EER has been manifold. They have: 1. Allowed for the development and testing of theory in a variety of ways; 2. Provided insight into the context specificity of circumstances and challenges of education systems, schools, teachers and students in particular countries; and 3. Advanced the methodology of the field. Each of these is now considered in turn.

First, critics of early EER drew attention to a perceived lack of theory that underpinned studies of educational effectiveness (Scheerens, 2016). Over time, EER theory has developed to better frame the complex underlying realities of schools, classrooms and education systems (Creemers, Kyriakides, & Sammons, 2010a). The dynamic model of educational effectiveness (Creemers & Kyriakides, 2008) has been particularly influential, accounting for the factors at the classroom, school, and system level – and interactions between these

– theorised to be associated with student outcomes (see Chapter 3; Kyriakides, Creemers & Panayiotou, n.d.). Scheerens's (2015) conceptualisation of educational *ineffectiveness* has also constituted an important contribution to theory by drawing attention to factors associated with negative student outcomes and/or small effect sizes. Studies in the context of countries where EER took hold early on have often been responsible for contributing to these developments in EER theory, and have extended it to take into account factors and organisational structures beyond the scope of schools and classrooms. As Muijs (n.d.) describes in Chapter 5 of this volume, for example, researchers in the UK and USA have gone beyond the proto-typical early EER considerations of school and classroom levels to investigate effectiveness, equity and improvement in education systems/countries (e.g. Van Damme, Liu, Vanhee, & Pustjens, 2010), districts/local authorities (e.g. Caldas & Bankston, 1999; Tymms et al., 2008), and networks of schools (e.g. Ainscow, Dyson, Goldrick & West, 2012; Chapman & Muijs, 2014; Lindorff, 2016; Muijs, Ainscow, Chapman, & West, 2011).

Second, the proliferation of single-country studies that take place in low-income countries and in countries with a more recent history of EER has contributed insights into the variation of school and teacher factors that account for differences in student outcomes, the variation (and sometimes consistency) of challenges faced by school, teachers, educational networks and systems, and the extent to which these factors and challenges “travel” across countries. For example, Creemers (1999) found that the classroom-level factors linked with student learning outcomes in Indonesian schools were as expected based on the EER knowledge generated in higher-income Western nations. However, the factors at the system level that affected practices within these classrooms were not necessarily accounted for by this prior research carried out elsewhere. Furthermore, some studies outside of established EER national settings evidence a more dramatic reconceptualization of definitions and relationships; the work of Harrison and Kuint (1998) in Israel, for example, demonstrated how effectiveness was reframed in terms of effective school responses to emergency in Israel. In other words, as researchers in a widening variety of locations began to engage with EER and to conduct EER, this did far more than provide specific empirical insights from individual

countries; it also afforded the opportunity to test, develop and extend theory, and it did so while raising new and important questions about context and the definition of “effectiveness”.

Third, a substantial body methodological literature in EER had been produced by the end of the 20<sup>th</sup> century with this methodology advanced so as to enable researchers to answer increasingly complex questions concerning the who, what, why, when, where, and how of educational effectiveness (Creemers, Kyriakides, & Sammons, 2010b). However, the extent to which the EER that takes place outside of countries in which EER was earliest established has capitalised on these methodological advances has varied. Multiple reasons for this have been suggested, but researcher capacity and training may be two key explanations. For example, Murillo (2007), in his account of how the EER literature has developed in Latin America since the first studies in the 1970s, notes the proliferation of “more and better studies, greater awareness of the specialized literature, and a new generation of well-trained researchers” (p. 86), and calls attention to the increasing use of sophisticated research methods (e.g. multilevel models, appropriate controls for student background) in that region. Recent examples include innovative EER designs investigating school and teacher effects in Chile (see for example Ortega, Malmberg, & Sammons, 2018a; 2018b). By contrast, it has been suggested that EER studies in Africa were for a long time limited in both number and methodological sophistication by the relative dearth of trained indigenous researchers who could undertake such studies (Fleisch, 2007), alongside practical difficulties in both appropriate measurement and data collection. The researcher capacity issue has begun to be addressed in some countries (see Chapter 8 of this volume by Bambanota et al., n.d., for an example from the Democratic Republic of the Congo), but the importance of involving researchers with local knowledge and understanding of education systems in low-income countries is an area that still requires attention and effort in the interest of a more inclusive, informed and context-sensitive international EER knowledge base.

It is worth noting that although the vast majority of EER has consisted of single-country studies, there have also been numerous collaborative efforts that have brought together and discussed collections of studies from around the world in books and in special issues of

peer-reviewed journals. Some of these collections, like those containing collected ICSEI country reports noted above, have specifically focused on scoping the field with an emphasis on international perspectives. Others have contributed in a more implicit way to international perspectives in the field by gathering researchers together from around the globe to share insights on a particular topic (for an example in teacher effectiveness, see Charalambous & Praetorius, 2018; in early childhood research, see Sammons, Anders, & Hall, 2013; and in methodology, see Sammons & Luyten, 2009).

Of course, there are limitations to the insights that can be gleaned from studies conducted in single countries. For example, the variations in local capacity, methodology and measurement noted above, plus differences in the languages of publication, have led to an unbalanced knowledge base in favour of English-language publications and studies conducted in higher-income countries. Additionally, although single-country studies can generate useful understandings of contexts, a comparative perspective is necessary to more fully understand the interactions between contextual and cultural characteristics and features of education systems, schools and classrooms, this being an important line of enquiry for future research in EER (Reynolds, et al., 2016).

## **2.5 The contribution of literature reviews and meta-analyses as syntheses of evidence**

As the number of EER studies has increased, and with the increasing diversity of countries in which EER is taking place, there is a growing contribution towards international perspectives in EER beyond that of single country studies. Reviews and meta-analyses of these studies are increasingly being carried out and these seek to synthesise findings across contexts and (sometimes) countries in order to advance the EER knowledge base. On the most basic level, such syntheses of evidence facilitate investigations of which system, network, school, teacher and classroom factors are most strongly related to student outcomes, and the extent to which variation at each level of EER (see Chapter 14; Hall, Lindorff, & Sammons, n.d.) is consistent across

different settings. Beyond that, these meta-analyses and reviews have also allowed for the extension and testing of theory in EER and have raised important questions about context-specificity.

Existing literature reviews provide information about the range of countries in which EER has been conducted, and offer insight regarding some of the differences and similarities between these countries in terms of the research methodologies used, empirical findings, and contextual and cultural factors. Although early reviews reflected the above-noted “ethnocentricity” (see Section 2) n generally referring to studies within the authors’ own national context(s), examples from the first *International Handbook of School Effectiveness Research* (Teddlie & Reynolds, 2000) onwards have more commonly explicitly attended to the international aspect and considered patterns, for example, across low- and high-income countries (e.g. Scheerens, 2001; Thomas, Kyriakides, & Townsend, 2016). Recent state-of-the-art narrative reviews of the literature in educational effectiveness (Reynolds et al., 2014), teacher effectiveness (Muijs et al., 2014), and school and system improvement (Hopkins, Stringfield, Harris, Stoll, & Mackay, 2014) have all to some extent noted the need for, and ways of, dealing with context-specificity, including what Thomas et al. (2016) term “micro aspects of context” (i.e. context at the level of the school and of the student; p. 220). Reviews focusing on specific regions or combinations of countries have also usefully provided syntheses of localised EER findings and foci, both in settings in which EER has had a relatively long tradition (see Chapter 10 of this volume by Opendakker, n.d., for an example from Belgium and the Netherlands) and in regions where EER has been more recently adopted (e.g. see Yu, 2007, for an example on Sub-Saharan Africa).

While literature reviews can provide syntheses of the EER evidence in a narrative framework (e.g. the state-of-the-art reviews noted above) , meta-analyses directly compare effect sizes across EER studies to generate generalisations including -- but not limited to -- aspects such as the relative size of school and teacher effects, the variations of these, and the associations between system, school and teacher variables with student outcomes (e.g. Scheerens & Bosker, 1997). It is worth noting that the development of a multilevel approach to meta-analysis (Raudenbush & Bryk, 1985) has contributed

to the provision of more robust syntheses of the EER evidence base, just as multilevel modelling constituted an important advance in primary EER methodology. However and as noted above for literature reviews, not all meta-analyses have focused on or necessarily even acknowledged the locations within which each included study was conducted (e.g. Hattie, 2009; Seidel & Shavelson, 2007), but there have been several that consider country of origin. For example, a fairly recent meta-analysis by Scheerens, Witziers, and Steen (2013) did this explicitly via a country moderator variable included in their multilevel analysis, although partially (the countries included the USA, the Netherlands, and “other countries”; p.627), and found some significant differences between countries in the effects of variables including parental involvement (with higher effect sizes in the USA than elsewhere) and curriculum quality (with lower effect sizes in the Netherlands than elsewhere). Furthermore, Kyriakides, Christoforou and Charalambous (2013), who used meta-analysis as a means of validating factors at the teacher level within the dynamic model of educational effectiveness, similarly included a country/area variable (including the USA, Europe, Asia, and other countries) though they found that this did not significantly predict effect sizes. This was interpreted to suggest that the teaching factors of the dynamic model seem to be generic, at least across the countries/areas included in that particular analysis. Common across both these examples, though, is the mixture of countries with groups of countries. This speaks to the disproportional weighting of existing EER research across countries and the continuing need to develop the international perspective in EER. In particular, there is scope for further research to investigate associations between effect sizes and finer-grained country variables – across a wider range of countries, perhaps – within a multilevel meta-analytic approach.

Of course, there are some inherent challenges and limitations from synthesising empirical findings, be this via narrative review or statistical meta-analysis. The methods used in individual studies vary, often making it difficult to untangle with any certainty that differences in findings might have resulted from underlying differences across countries rather than from differences in methodology. Further, depending on the approach taken to synthesising evidence across multiple studies, there is a risk of overlooking the importance of context.

Avoiding this requires attention to the challenges encountered, methodological approaches, and limitations of the studies considered, as well as “effects” at each level (e.g. school, classroom, pupil) and of individual variables, to make sense of any differences across contexts.

## **2.6 The contribution of international studies within and beyond EER: Evidence on cross-country comparisons**

Despite the above noted dominance of single-country studies within the EER literature, comparative international studies focusing on student outcomes – particularly educational achievement – in order to evaluate the equity and quality of education systems have existed for approximately same period (about half a century) as have studies of school effectiveness (see Section 2). International large-scale assessments (ILSAs) began in the 1960s, with a study of mathematics achievement in twelve countries led by the International Association for the Evaluation of Educational Achievement (IEA; Foshay, Thorndike, Hotyaat, Pidgeon, & Walker, 1962).

In the five decades that followed the first IEA twelve-country study, ILSAs have gained increasing prominence though why this is the case is likely due to multiple factors. Plausible reasons include the international policy trend towards globalisation, ILSA’s contributions to that trend, or a combination of the two. Either way, the number of countries participating in ILSAs has increased over time, from twelve countries in the 1962 IEA mathematics study to 80 jurisdictions in the 2018 administration of the Programme for International Assessment (PISA; OECD, 2018). Results of country comparisons have consequently drawn considerable attention from policymakers. There is considerable evidence to suggest that ILSAs and associated surveys have informed policy changes in a number of countries, particularly with regards to curriculum reform and performance standards (Lietz & Tobin, 2016), although in some cases reforms using ILSA results as a rationale have been based on misinterpretations or policy-maker cherry-picking (see Klieme, n.d., in Chapter 7 of this volume). The Trends in International Mathematics and Science Study (TIMSS; first conducted in 1995) and Progress in International

Reading Literacy Study (PIRLS; first administered in 2001) conducted by the IEA, and the PISA (conducted every three years since 2000) and Teaching and Learning International Survey (TALIS; first administered in 2008) conducted by the Organisation for Economic Co-operation and Development (OECD) have been particularly influential, with country league table rankings widely publicised in the mass media.

Critics of these studies have pointed out a number of methodological and practical limitations. Reynolds et al. (n.d.) summarise these in detail in Chapter 6 of this volume – though noting that not all of these limitations apply to each ILSA – pointing to challenges in designing assessments with cross-cultural validity, sampling issues, often cross-sectional design, limited measurement of student background variables, limited measurement of teaching and learning processes, heavy emphasis on resource-based school factors, and limited measurement of non-academic student outcomes. Some of these limitations have been addressed by more recent developments; for example, the most recent PISA studies have adopted conceptual frameworks drawing on EER with input from established EER researchers (Klieme & Kuger, 2014; OECD, 2009), and a pilot of direct observation of video-recorded mathematics lessons to extend TALIS is currently underway at the time of writing this chapter (OECD, 2017) involving 8 countries/jurisdictions including China (Shanghai), Spain (Madrid), Chile, Colombia, Japan, Mexico, England, and Germany.

While ILSAs have been in some ways quite distinct from EER, and the educational effectiveness literature tended for some time not to refer to ILSA literature (Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002), in some respects there has been reciprocal influence and synergy between the two. As noted above, EER researchers (e.g. Eckhard Klieme and Jaap Scheerens) have acted as experts and advisors for the OECD in order to develop and extend conceptual and analytic frameworks, and have featured in teams conducting OECD country reviews that have drawn on EER and education improvement perspectives (see OECD, 2013, p. 663 for the list of country review team members; see Mortimore, Field, & Pont, 2004; Nusche, Laveault, MacBeath, & Santiago, 2011; Santiago, Gilmore, Nusche, & Sammons, 2012; Shewbridge, Ehren, Santiago, & Tamassia, 2012,

for examples of thematic reviews and reports). Conversely, EER researchers have also used ILSA data for secondary analyses to extend the EER knowledge base and to develop and test EER theory, dating back to re-analysis of IEA data to investigate the generalisability of effectiveness factors across countries (see Scheerens, Vermeulen, & Pelgrum, 1989, for an early example of this). More recent secondary analyses of ILSA data have provided new types of insights that are particularly pertinent to EER. For example, a secondary analysis of TIMSS data (e.g. Kyriakides, 2006b) indicated a smaller proportion of variance between students, suggesting larger higher-level effects than generally identified in national studies, and secondary analyses of PISA data have provided insight into the functioning of specific factors across countries (e.g. Ning, Van Damme, Van Den Noortgate, Yang, & Gielen, 2015, on classroom disciplinary climate) and introduced country-level effectiveness measures for assessing contextualised effectiveness rather than simply identifying high-performing systems (e.g. Lenkeit & Caro, 2014), to name a few.

Beyond the re-analysis of data from ILSAs, researchers situated within EER have also drawn attention to the importance of primary comparative international studies within the field in order to extend existing theory and understanding of the impact of culture and context on schools, teachers and students. Some of the key reasons suggested (e.g. by Kyriakides, 2006a; and Reynolds, 2000) for why international EER studies have an important contribution to make have included:

- Provision of an evidence base for (or against) the trend towards policy transfer across countries
- Correct and thorough identification of the power of school and classroom variation and effects of school and classroom variables beyond what is possible within single-country studies
- Generation of more sensitive theoretical explanations both to inform research design and to inform interpretation and understanding of results

Reynolds (2000) has further specified the need to study the relationships between processes, outcomes, and contexts in order to understand how different instructional variables relate to student outcomes in different contexts, and also how different school and education system features may foster similarly effective classroom practices in different contexts.

In response to this perceived need for comparative international studies within EER, one of the earliest examples of such work was the International School Effectiveness Research Project (ISERP; Reynolds, 2006; Reynolds et al., 2002). The aims of this study were to investigate “which factors are associated with student academic and social outcomes across countries and which factors are restricted to certain cultural context”, as well as “which factors are associated with student academic and social outcomes across countries, for students with different characteristics” (i.e. differential effectiveness within and across countries; Reynolds, 2006, p. 539). The mixed-method research design was innovative and comprehensive, including observations of classroom and school processes, affective and social outcome measures, and a longitudinal cohort design. However, the study also highlighted some of the fundamental challenges in conducting such comparative international studies. Cultural differences in the discourse surrounding educational effectiveness, together with the above-mentioned dominance of certain countries within the EER knowledge base, meant that ensuring the cultural relevance of educational factors was a challenge. Similarly, social outcomes were found to be more culturally-specific than academic ones, and therefore more difficult to measure and compare across regions and cultures; cross-cultural high-inference judgments of teaching practice were similarly found to be problematic. Nonetheless, the challenges and limitations of ISERP are arguably as (or more) useful towards an international perspective in EER as are the actual findings of the study; the detailed and transparent accounts of the research process (e.g. in Reynolds et al., 2002) have the potential to guide future comparative international EER efforts in addressing context specificity and cultural differences.

Within teacher effectiveness, the above-mentioned International System for Teacher Observation and Feedback (ISTOF; Muijs et al., 2018; Teddlie, Creemers, Kyriakides, Muijs, & Yu, 2006) was an ambitious effort to develop and validate an instrument for classroom observation across 20 participating countries (with at least some representation of regions including Europe, East Asia, South Asia, Southeast Asia, Africa, South America, and North America). The ISTOF instrument, developed using a modified Delphi technique drawing interactively on expert opinion and review, has since been validated and used in additional settings beyond the countries that

contributed to its development (e.g. Soderlund, Sorlie, & Syse, 2015) as well as used for additional comparative studies (see Miao, Reynolds, Harris, & Jones, 2015, for an example comparing teaching in England and China).

Comparative international research within the area of school improvement has been characterised as having a more international character earlier in its history than was the case for educational effectiveness (Reynolds, 2000), and has also fed into the designs of studies at the intersection of EER and school improvement. The International School Improvement Project (ISIP; Bollen & Hopkins, 1987) provides an example of an early school-improvement focussed project on an international scale (with 14 participating countries), sponsored by the OECD though notably in contrast to the ILSA paradigm. Although Hopkins (1990) proposed that ISIP's focus on improvement strategies, goal-setting and processes marked a clear contrast with the effective schools knowledge base at the time, the detailed contextual information gleaned from the study's in-depth approach had the potential to inform more context-sensitive methods and measures in EER. Another example, the International Successful School Principals Project (ISSPP), which built upon an approach used by Day and colleagues (2000) in the UK, was initiated in 2001 with researchers from seven countries and grew to include fourteen countries by 2010. This study explored the qualities, behaviours and contexts of successful school principals, with a broad conceptualisation of success based on multiple indicators including but not limited to evidence of student achievement. The ISSPP constituted a large-scale, case-study based investigation of the role of school leaders in school improvement, and generated insight within and across contexts (e.g. Crow, Day, & Møller, 2017; Gurr, 2014; Johnson, Møller, Jacobson, & Wong, 2008; also see Chapter 12 in this volume by Townsend, Berryman, Gurr, & Drysdale, n.d.); findings also provided information about the relationship between principal leadership and student learning (e.g. Day, Gu, & Sammons, 2016), which has significant implications for EER. The Effective School Improvement project (Creemers, 2002) comprised another important example of a comparative international study that advanced an international perspective, with an explicit link between the case study approach used and a set of guiding principles drawn from the existing EER knowledge base.

Replications of EER studies and improvement projects have been transferred across settings, too. For example, Professional Learning Communities based on previous research in the UK (Bolam, McMahon, Stoll, Thomas, & Wallace, 2005) are being implemented in China as part of the broader Improving Teacher Development and Educational Quality in China (ITDEQC) project (Thomas et al., 2016). On one hand, such examples of the uptake of, and action based on, findings from EER and school improvement initiatives from one setting to another are promising in that they have involved input from researchers based in both the original and the new contexts, allowing for local insight to inform implementation. On the other hand, the long-term success of such initiatives merits further investigation.

More recent examples of comparative international efforts in EER have included a mixture of different approaches, each with unique contributions to make towards the EER knowledge base. In addition to the projects mentioned above that included countries from different regions of the globe, regional multi-country studies help to shed light on within-region variation and inform the design of future studies (see Kelcey & Shen, 2016, for an example investigating school and teacher effects across 15 Sub-Saharan African countries). Two-country comparative studies have fewer practical complications to implement, yet still contribute insight into cultural and contextual differences in educational factors and student outcomes (see van de Grift, Chun, Maulana, Lee, & Helms-Lorenz, 2017, for an example within teacher effectiveness comparing across South Korea and the Netherlands).

Despite the developments within and beyond EER towards an international perspective via comparative international studies highlighted in this section, further work is needed to more fully understand how national or system-level policies and cultural factors affect schools, teaching and students (see Chapter 3 of this volume by Kyriakides, Creemers, & Panayiotou, n.d.) and thus support or hinder their effectiveness in promoting different student outcomes. Lessons learned from the successes and also from the limitations of past studies, as well as across the areas of EER and ILSA, have the potential to inform increasingly rigorous and comprehensive future comparative international research in EER. This in turn has the potential to

inform policy and practice across the system, school and classroom levels in multiple contexts.

## **2.7 Reflections and suggestions for future directions**

There is no question that EER has become an increasingly international field since the earliest studies conducted in the 1970s. Recent EER studies in middle- and low-income countries have begun to demonstrate the use of more sophisticated and up-to-date research methods and models (both theoretical and statistical), and a wider variety of settings have been considered to contribute to the international evidence base. International comparative studies have become more common, perhaps at least in part because of increased synergy between EER and ILSA, and researchers seem to have generally become more outward looking in their reviews of the literature to contextualise single-country empirical studies.

EER researchers have also begun to employ or at least call for a greater diversity of approaches to research design and analysis. The increasing convergence and mutual learning between educational effectiveness and school improvement research has afforded opportunities to better account for and attend to context. Mixed methods studies (e.g. ISERP) have demonstrated how the combination of qualitative and quantitative approaches can provide more thorough and robust insight into contextual understandings of concepts and educational processes. New statistical approaches to investigating educational effects (e.g. Sammons & Luyten, 2009) and consideration of small-scale experimental studies to test innovations in particular school contexts (Muijs et al., 2014) have the potential to contribute to the international knowledge base while maintaining attention to context specificity and cultural features. If the purpose of EER is ultimately to inform not only more nuanced understandings of educational quality and equity (Kyriakides, Creemers, & Charalambous, 2018) but also to support the development of more effective and equitable education (Chapman et al., 2012), such insight is essential.

However, there are still a number of issues that we suggest are in need of further investigation with regard to international perspectives

in EER. Critical changes in national and international politics have taken place within the last decade, and the implications of these for the international landscape of EER, not to mention broader consequences for education in general and conceptualisations of what and for whom schooling is meant to function, have yet to be seen. We live in an increasingly globalised world, but many localised problems persist while others emerge and pose unique challenges (e.g. in response to natural disasters and violent conflict). We know a great deal about educational effectiveness, but much still remains to be learned, and many problems remain to be solved.

Reflecting on the above narrative account of international perspectives in the field, then, and on the political and policy contexts in which this narrative is contextualised, future work is needed to develop and assess the state of the field with regard to international perspectives in four areas:

1. Previous reviews of the literature have drawn attention to themes and developments from research in “developing” or “emerging” contexts, but a comprehensive, up-to-date systematic review of the international literature in EER is due.
2. Long-term research is needed to provide robust evidence of the uptake and transfer of educational innovations and initiatives between diverse contexts, in order to inform deeper understandings of why, how and where between-country or between-culture practices and policies “travel” (or conversely fail to “travel”).
3. More research is still needed to develop a meaningful account that explains how concepts and educational factors are perceived and operationalised in different contexts, something Reynolds et al. (2016) have raised an “a matter for urgent attention” (p. 412).
4. On-going capacity building remains an essential consideration in places where EER research has generally been conducted by external researchers, to ensure that EER is informed by appropriate local knowledge and understanding.

International perspectives in EER can inform, and should be informed by, critical debates “about what education is for, about how it should be organized and – indeed – about the kind of society in which we wish to live” (to echo the words of Mortimore, 2013, p. 236). Further,

we agree that education is ‘under siege’ and that educational research – including EER – plays a critical role within a democratic discourse on education policy,

“...Research findings frequently challenge the expectations of policymakers. This discord, awkward though it may be, is essential in a democracy. The 1950s Quaker adage applies aptly to educational research when it states the need “to speak truth to power”” (Mortimore, 2010).

Managing productive relationships and facilitating dialogue across stakeholder groups (policymakers, practitioners and researchers) is an important part of that democratic discourse (see Chapman, Ainscow & Hadfield, 2020).

Ultimately, the development of an internationally diverse body of EER knowledge can advance our understanding of education (from the system-level to the level of the classroom), it can inform evidence-based and appropriately contextualised educational policy, and it can inform meaningful educational improvements and reform. To help inform progress towards these goals, this chapter has provided a historic overview of the development of this body of knowledge to date. This helps signpost future directions of EER and it sets the scene for the other chapters in this volume.

## References

- Ainscow, M., Dyson, A., Goldrick, S. & West, M. (2012). *Developing equitable education systems*. London: Routledge.
- Bambanota, G., Van Damme, J., De Fraine, B., Masimango, P. V., Mabela, G. K., Tshite, A. M., ... Legono, J. P. B. (n.d.). Educational effectiveness research in Africa: The case of the Democratic Republic of the Congo (DRC). In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Bolam, R., McMahon, A., Stoll, L., Thomas, S., & Wallace, M. (2005). *Creating and sustaining professional learning communities*. Bristol: University of Bristol.
- Bollen, R., & Hopkins, D. (1987). *School based research: Towards a praxis*. Leuven [Belgium]: ACCO Publishing.
- Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K., & Wisenbaker, J. M. (1978). Elementary School Social Climate and School Achievement. *American Educational Research Journal*, 15(2), 301–318.
- Caldas, S. J., & Bankston, C. L. (1999). Multilevel Examination of Student, School, and District-Level Effects on Academic Achievement. *The Journal of Educational Research*, 93(2), 91–100. <https://doi.org/10.1080/00220679909597633>

- Chapman, C., Ainscow, M. & Hadfield, M. (2020). *Changing education systems: A research based approach*. London: Routledge.
- Chapman, C., Armstrong, P., Harris, A., Muijs, D., Reynolds, D., & Sammons, P. (2012). *School effectiveness and improvement research, policy, and practice challenging the orthodoxy?* London; New York: Routledge.
- Chapman, C., & Muijs, D. (2014). Does school-to-school collaboration promote school improvement? A study of the impact of school federations on student outcomes. *School Effectiveness and School Improvement*, 25(3), 351–393.
- Chapman, C., Reynolds, D., Muijs, D., Sammons, P., Stringfield, S., & Teddlie, C. (2016). Educational effectiveness and improvement research and practice: The emergence of the discipline. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *The Routledge international handbook of educational effectiveness and improvement* (pp. 1–24). London: Routledge.
- Charalambous, C. Y., & Praetorius, A.-K. (2018). Studying instructional quality in mathematics through different lenses: in search of common ground. *ZDM: The International Journal on Mathematics Education*, 50(3).
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, D.C.: US Department of Health, Education, and Welfare.
- Creemers, B. (1994). *The effective classroom*. London: Cassell.
- Creemers, B. (1999). Educational effectiveness and improvement in developing societies: Some experiences from the Primary Education Quality Improvement Project in Indonesia. *Tertium Comparationis*, 5(1), 32–51.
- Creemers, B. (2002). From School Effectiveness and School Improvement to Effective School Improvement: Background, Theoretical Analysis, and Outline of the Empirical Study. *Educational Research and Evaluation*, 8(4), 343–362. <https://doi.org/10.1076/edre.8.4.343.8814>
- Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness : a contribution to policy, practice and theory in contemporary schools*. London; New York: Routledge.
- Creemers, B., Kyriakides, L., & Sammons, P. (2010a). *Methodological Advances in Educational Effectiveness Research*. Milton Park, Abingdon, Oxon; New York: Routledge.
- Creemers, B., Kyriakides, L., & Sammons, P. (2010b). *Methodological Advances in Educational Effectiveness Research*. Milton Park, Abingdon, Oxon; New York: Routledge.
- Creemers, B., Peters, T., & Reynolds, D. (1989). *School effectiveness and school improvement*. Amsterdam: Swets & Zeitlinger.
- Creemers, B., Reynolds, D., & Swint, F. E. (1996). *The International School Effectiveness Research Project quantitative and class study data*.
- Crow, G., Day, C., & Møller, J. (2017). Framing research on school principals' identities. *International Journal of Leadership in Education*, 20(3), 265–277. <https://doi.org/10.1080/13603124.2015.1123299>
- Day, C., Gu, Q., & Sammons, P. (2016). The Impact of Leadership on Student Outcomes: How Successful School Leaders Use Transformational and Instructional Strategies to Make a Difference. *Educational Administration Quarterly*, 52(2), 221–258. <https://doi.org/10.1177/0013161X15616863>
- Day, C., Harris, A., Hadfield, M., Tolley, H., & Beresford, J. (2000). *Leading schools in times of change*. Maidenhead: Open University Press.
- Doyle, W. (1977). Paradigms for Research on Teacher Effectiveness. *Review of Research in Education*, 5(1), 163–198. <https://doi.org/10.3102/0091732X005001163>
- Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15–27.
- Education Endowment Foundation. (2018). Sutton Trust-EEF Teaching and learning toolkit & EEF early years toolkit: Technical appendix and process manual. London: Education Endowment Foundation.

- Fleisch, B. (2007). History of the school effectiveness and improvement movement in Africa. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 341–350). Dordrecht: Springer.
- Foshay, A. W., Thorndike, R. L., Hotyaat, F., Pidgeon, D. A., & Walker, D. A. (Eds.). (1962). *Educational achievements of thirteen-year-olds in twelve countries*. Hamburg: UNESCO Institute for Education.
- Gage, N. L. (1963). *Handbook of research on teaching*. (N. L. Gage, Ed.). Chicago: Rand McNally.
- Gurr, D. (2014). Successful school leadership across contexts and cultures. *Leading and Managing*, 20(2), 75–88.
- Hall, J., Lindorff, A., & Sammons, P. (2016). *Evaluation of the Impact and Implementation of Inspire Maths in Year 1 Classrooms in England*. Oxford: University of Oxford. <https://doi.org/10.13140/RG.2.2.23042.07369>
- Harris, A., & Chrispeels, J. (2006). *Improving schools and educational systems: international perspectives*. London; New York: Routledge.
- Harrison, J., & Kuint, S. (1998). School Culture and School Effectiveness in Emergencies: Lessons from Israeli Experience during the Gulf War. *School Effectiveness and School Improvement*, 9(2), 192–217. <https://doi.org/10.1080/0924345980090204>
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge. <https://doi.org/10.4324/9780203887332>
- Hopkins, D. (1990). The International School Improvement Project (ISIP) and Effective Schooling: towards a synthesis. *School Organisation*, 10(2–3), 179–194. <https://doi.org/10.1080/0260136900100204>
- Hopkins, D., Stringfield, S., Harris, A., Stoll, L., & Mackay, T. (2014). School and system improvement: a narrative state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 257–281. <https://doi.org/10.1080/09243453.2014.885452>
- Jaciw, A. P., Hegseth, W. M., Lin, L., Toby, M., Newman, D., Ma, B., & Zacamy, J. (2016). Assessing Impacts of Math in Focus, a “Singapore Math” Program. *Journal of Research on Educational Effectiveness*, 9(4), 473–502. <https://doi.org/10.1080/19345747.2016.1164777>
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., ... Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jerrim, J., & Vignoles, A. (2015). *The causal effect of East Asian “mastery” teaching methods on English children’s mathematical skills?* (No. 15–05). London.
- Johnson, L., Møller, J., Jacobson, S. L., & Wong, K. C. (2008). Cross-national Comparisons in the International Successful School Principalship Project (ISSPP): The USA, Norway and China. *Scandinavian Journal of Educational Research*, 52(4), 407–422. <https://doi.org/10.1080/00313830802184582>
- Kelcey, B., & Shen, Z. (2016). Multilevel design of school effectiveness studies in sub-Saharan Africa. *School Effectiveness and School Improvement*, 27(4), 492–510. <https://doi.org/10.1080/09243453.2016.1168855>
- Klieme, E. (n.d.). Policies and practices of assessment: A showcase for the use (and misuse) of International Large Scale Assessments in educational effectiveness research. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Klieme, E., & Kuger, S. (2014). PISA 2015 draft questionnaire framework. Paris: OECD. Retrieved from <http://www.oecd.org/pisa/pisaproducts/PISA-2015-draft-questionnaire-framework.pdf>
- Kratz, H. E. (1896). Characteristics of the Best Teacher as Recognized by Children. *The Pedagogical Seminary*, 3(3), 413–460. <https://doi.org/10.1080/08919402.1896.9943604>
- Kyriakides, L. (2006a). Introduction international studies on educational effectiveness. *Educational Research and Evaluation*, 12(6), 489–497. <https://doi.org/10.1080/13803610600873960>

- Kyriakides, L. (2006b). Using international comparative studies to develop the theoretical framework of educational effectiveness research: A secondary analysis of TIMSS 1999 data. *Educational Research and Evaluation*, 12(6), 513–534. <https://doi.org/10.1080/13803610600873986>
- Kyriakides, L., Christoforou, C., & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, 36, 143–152. <https://doi.org/https://doi.org/10.1016/j.tate.2013.07.010>
- Kyriakides, L., Creemers, B., & Panayiotou, A. (n.d.). Developing and testing theories of educational effectiveness addressing the dynamic nature of education. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Kyriakides, L., Creemers, B., & Charalambous, E. (2018). *Equity and quality dimensions in educational effectiveness*. Cham, Switzerland: Springer.
- Lenkeit, J., & Caro, D. H. (2014). Performance status and change – measuring education system effectiveness with data from PISA 2000–2009. *Educational Research and Evaluation*, 20(2), 146–174. <https://doi.org/10.1080/13803611.2014.891462>
- Lietz, P., & Tobin, M. (2016). The impact of large-scale assessments in education on education policy: evidence from around the world. *Research Papers in Education*, 31(5), 499–501. <https://doi.org/10.1080/02671522.2016.1225918>
- Lindorff, A. (2016). *Networks for school support and improvement: A mixed methods study*. University of Oxford.
- Lindorff, A., Hall, J., & Sammons, P. (2019). Investigating a Singapore-Based Mathematics Textbook and Teaching Approach in Classrooms in England. *Frontiers in Education*, 4(37). <https://doi.org/10.3389/educ.2019.00037>
- Miao, Z., Reynolds, D., Harris, A., & Jones, M. (2015). Comparing performance: a cross-national investigation into the teaching of mathematics in primary classrooms in England and China. *Asia Pacific Journal of Education*, 35(3), 392–403. <https://doi.org/10.1080/02188791.2015.1056593>
- Mortimore, P. (1998). *The road to improvement: Reflections on school effectiveness*. Abington: Taylor & Francis.
- Mortimore, P. (2001). Globalisation, Effectiveness and Improvement. *School Effectiveness and School Improvement*, 12(2), 229–249. <https://doi.org/10.1076/sesi.12.2.229.3453>
- Mortimore, P. (2010, June 1). The education ministers must base their policies on sound research. *The Guardian*. Retrieved from <https://www.theguardian.com/education/2010/jun/01/education-research-sound-policy-base>
- Mortimore, P. (2013). *Education under siege: Why there is a better alternative*. Bristol: Policy Press.
- Mortimore, P., Field, S., & Pont, B. (2004). *Equity in education: Thematic review*. Paris: OECD.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). *School matters: The junior years*. Somerset: Open Books.
- Muijs, (n.d.). Extending educational effectiveness: The middle tier and network effectiveness. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Muijs, D., Ainscow, M., Chapman, C., & West, M. (2011). *Collaboration and Networking in Education*. Dordrecht: Springer Science+Business Media B.V.
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., & Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P. M., & Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: how useful is the International System for Teacher Observation and Feedback (ISTOF)? *ZDM*, 50(3), 395–406. <https://doi.org/10.1007/s11858-018-0921-9>

- Murillo, J. (2007). School effectiveness research in Latin America. In T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 75–91). Dordrecht: Springer.
- Ning, B., Van Damme, J., Van Den Noortgate, W., Yang, X., & Gielen, S. (2015). The influence of classroom disciplinary climate of schools on reading achievement: a cross-country comparative study. *School Effectiveness and School Improvement*, 26(4), 586–611. <https://doi.org/10.1080/09243453.2015.1025796>
- Nusche, D., Laveault, D., MacBeath, J., & Santiago, P. (2011). *OECD Reviews of evaluation and assessment in education: New Zealand 2011*. Paris: OECD.
- OECD. (2009). PISA 2009 assessment framework: Key competencies in reading, mathematics and science. Paris: OECD. Retrieved from <https://www.oecd.org/pisa/pisaproducts/44455820.pdf>
- OECD. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. Paris: OECD.
- OECD. (2017). TALIS 2018: Video study and global video library on teaching practices. Retrieved October 30, 2018, from <http://www.oecd.org/education/school/TALIS-2018-video-study-brochure-ENG.pdf>
- OECD. (2018). About - PISA. Retrieved March 1, 2019, from <https://www.oecd.org/pisa/aboutpisa/pisa-2018-participants.htm>
- Opdenakker, M.-C. (n.d.). Three decades of educational effectiveness research in Belgium and the Netherlands: Key studies, main research topics and findings. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Ortega, L., Malmberg, L.-E., & Sammons, P. (2018a). School effects on Chilean children's achievement growth in language and mathematics: An accelerated growth curve model. *School Effectiveness and School Improvement*, 29(2), 308–337. <https://doi.org/10.1080/09243453.2018.1443945>
- Ortega, L., Malmberg, L.-E., & Sammons, P. (2018b). Teacher Effects on Chilean Children's Achievement Growth: A Cross-Classified Multiple Membership Accelerated Growth Curve Model. *Educational Evaluation and Policy Analysis*, 40(3), 473–501. <https://doi.org/10.3102/0162373718781960>
- Plowden, B., & Central Advisory Council for England. (1967). *Children and their primary schools*. London.
- Raudenbush, S. W., & Bryk, A. S. (1985). Empirical Bayes Meta-Analysis. *Journal of Educational Statistics*, 10(2), 75–98. <https://doi.org/10.3102/10769986010002075>
- Reynolds, D. (2000). School effectiveness: The international dimension. In C. Teddlie & D. Reynolds (Eds.), *International handbook of school effectiveness research* (pp. 232–256). London: Falmer.
- Reynolds, D. (1996). Introduction and Overview. *School Effectiveness and School Improvement*, 7(2), 111–113. <https://doi.org/10.1080/0924345960070201>
- Reynolds, D. (2006). World Class Schools: Some methodological and substantive findings and implications of the International School Effectiveness Research Project (ISERP). *Educational Research and Evaluation*, 12(6), 535–560. <https://doi.org/10.1080/13803610600874026>
- Reynolds, D., & Bellin, W. (2002). *World class schools: international perspectives on school effectiveness*. London; New York: RoutledgeFalmer.
- Reynolds, D., Caldwell, B., Cruz, R. M., Miao, Z., Murillo, J., Mugendawala, H., Mayol, B.I., Medina, C.P. & Ramon, M. R. R. (2016). Comparative educational research. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *The Routledge international handbook of educational effectiveness and improvement* (pp. 246–282). London: Routledge.
- Reynolds, D., Chapman, C., Clarke, P., Muijs, D., Sammons, P., & Teddlie, C. (2016). The future of educational effectiveness and improvement research, and some suggestions and speculations. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *The Routledge international handbook of educational effectiveness and improvement* (pp. 408–439). London: Routledge.

- Reynolds, D., Creemers, B. P. M., Stringfield, S., Teddlie, C., & Schaffer, G. (2002). *World class schools : international perspectives on school effectiveness*. London: Routledge.
- Reynolds, D., Kelly, A., Harris, A., Jones, M., Adams, D., Miao, Z., & Bokhove, C. (n.d.). Extending educational effectiveness: A critical review of research approaches in international effectiveness research, and proposals to improve them. In J. Hall, A. Lindorff, & P. Sammons (Eds.), *International perspectives in educational effectiveness research*. Springer.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 197–230. <https://doi.org/10.1080/09243453.2014.885450>
- Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: secondary schools and their effects on children*. Cambridge, Mass.: Harvard University Press.
- Sammons, P., Anders, Y., & Hall, J. (2013). Educational effectiveness approaches in early childhood research across Europe. *School Effectiveness and School Improvement*, 24(2), 131–137. <https://doi.org/10.1080/09243453.2012.749790>
- Sammons, P., Davis, S., & Gray, J. (2016). Methodological properties of school effectiveness research: Exploring the underpinnings, evolution and future directions of the field. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *Routledge international handbook of educational effectiveness and improvement* (pp. 25–76). London: Routledge.
- Sammons, P., & Luyten, H. (2009). Editorial article for special issue on alternative methods for assessing school effects and schooling effects. *School Effectiveness and School Improvement*, 20(2), 133–143. <https://doi.org/10.1080/09243450902879753>
- Santiago, P., Gilmore, A., Nusche, D., & Sammons, P. (2012). *OECD Reviews of evaluation and assessment in education: Czech Republic 2012*. Paris: OECD.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford, OX; New York, N.Y.: Pergamon.
- Scheerens, Jaap. (2001). Monitoring School Effectiveness in Developing Countries. *School Effectiveness and School Improvement*, 12(4), 359–384. <https://doi.org/10.1076/sesi.12.4.359.3447>
- Scheerens, Jaap. (2015). Theories on educational effectiveness and ineffectiveness. *School Effectiveness and School Improvement*, 26(1), 10–31. <https://doi.org/10.1080/09243453.2013.858754>
- Scheerens, Jaap. (2016). Educational Effectiveness and Ineffectiveness a Critical Review of the Knowledge Base. Dordrecht: Springer Netherlands; Retrieved from <http://dx.doi.org/10.1007/978-94-017-7459-8>
- Scheerens, Jaap, Vermeulen, C. J. A. J., & Pelgrum, W. J. (1989). Generalizability of instructional and school effectiveness indicators across nations. *International Journal of Educational Research*, 13(7), 789–799. [https://doi.org/https://doi.org/10.1016/0883-0355\(89\)90029-3](https://doi.org/https://doi.org/10.1016/0883-0355(89)90029-3)
- Scheerens, Jaap, Witziers, B., & Steen, R. (2013). A meta-analysis of school effectiveness studies. *Revista de Educacion*, 361, 619–645.
- Schildkamp, K. (2019). President Welcome. Retrieved March 15, 2019, from <https://www.icsei.net/president-welcome/>
- Seidel, T., & Shavelson, R. J. (2007). Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Shewbridge, C., Ehren, M., Santiago, P., & Tamassia, C. (2012). *OECD Reviews of evaluation and assessment in education: Luxembourg 2012*. Paris: OECD.
- Society for Research on Educational Effectiveness (2019). *Membership*. Retrieved from <https://www.sree.org/membership>
- Soderlund, G., Sorlie, K., & Syse, I. (2015). Mestringsforventninger i matematikk. In *Finnut, Lilyhammer*.

- Stevens, R. (1912). *The question as a measure of efficiency in instruction*. New York: Teachers College, Columbia University. Retrieved from <https://archive.org/details/questionasmeasur00stevuoft>
- Teddlie, C., Creemers, B., Kyriakides, L., Muijs, D., & Yu, F. (2006). The international system for Teacher Observation and Feedback: Evolution of an international study of teacher effectiveness constructs 1. *Educational Research and Evaluation, 12*(6), 561–582.
- Teddlie, C., & Reynolds, D. (Eds.). (2000). *International handbook of school effectiveness research*. London: Falmer.
- Thomas, S., Kyriakides, L., & Townsend, T. (2016). Educational effectiveness in new, emerging, and traditional contexts. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons, & C. Teddlie (Eds.), *The Routledge international handbook of educational effectiveness and improvement* (pp. 220–245). London: Routledge.
- Townsend, T., Berryman, M., Gurr, D., & Drysdale, L. (n.d.). Leadership for learning in diverse settings: School leaders setting the agenda in Australia and New Zealand. In *International perspectives in educational effectiveness research*. Dordrecht: Springer.
- Townsend, T., Clarke, P., & Ainscow, M. (1999). *Third millennium schools : a world of difference in effectiveness and improvement*. Lisse, Netherlands; Exton, PA: Swets & Zeitlinger.
- Tymms, P., Merrell, C., Heron, T., Jones, P., Albone, S., & Henderson, B. (2008). The importance of districts. *School Effectiveness and School Improvement, 19*(3), 261–274. <https://doi.org/10.1080/09243450802332069>
- Van Damme, J., Liu, H., Vanhee, L., & Pustjens, H. (2010). Longitudinal studies at the country level as a new approach to educational effectiveness: explaining change in reading achievement (PIRLS) by change in age, socio-economic status and class size. *Effective Education, 2*(1), 53–84. <https://doi.org/10.1080/19415531003616888>
- van de Grift, W. J. C. M., Chun, S., Maulana, R., Lee, O., & Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *School Effectiveness and School Improvement, 28*(3), 337–349. <https://doi.org/10.1080/09243453.2016.1263215>
- Weber, G. (1971). *Inner city children can be taught to read: Four successful schools*. Washington, D.C.: Council for Basic Education.
- What Works Clearinghouse. (2017). *What Works Clearinghouse Standards handbook (Version 4.0)*. Washington, D.C.: National Center for Education Evaluation and Regional Assistance.
- Yu, G. (2007). *Research evidence of school effectiveness in Sub-Saharan Africa*. Bristol: EdQual Research Programme Consortium.