Chapter 3
School Improvement Capacity – A Review and a Reconceptualization from the Perspectives of Educational Effectiveness and Educational Policy

David Reynolds and Annemarie Neeleman

3.1 Introduction

The field of school improvement (SI) has developed rapidly over the last 30 years, moving from the initial Organisational Development (OD) tradition to school-based review, action research models, and the more recent commitment to leadership-generated improvement by way of instructional (currently) and distributed (historically) varieties. However, it has become clear from the findings in the field of educational effectiveness (EE) (Chapman et al., 2012; Reynolds et al., 2014) that SI needs to be aware of the following developmental needs based on insights from both EE (Chapman et al., 2012; Reynolds et al., 2014) and educational practice as well as other research disciplines, if it will be considered an agenda-setting topic for practitioners and educational systems.

3.1.1 What Kind of School Improvement?

Following Scheerens (2016), we interpret school improvement as the "dynamic application of research results" that should follow the research activity of educational effectiveness. Basically, it is the schools and educational systems that have been carrying out school improvement themselves over the years. However, this is poorly understood, rarely conceptualised/measured and, what is even more remarkable, seldom used as the design foundations of conventionally described SI. Many

D. Reynolds (✉)
Swansea University, Swansea, UK
e-mail: david@davidreynoldsconsulting.com; david.reynolds@swansea.ac.uk

A. Neeleman
Maastricht University, Maastricht, The Netherlands

policy-makers and educational researchers tend to cling to the assumption that EE, supported by statistically endorsed effectiveness-enhancing factors, should set the SI agenda (e.g. Creemers & Kyriakides, 2009). However logical this assumption may sound, educational practice has not been necessarily predisposed to act accordingly.

A recent comparison (Neeleman, 2019a) between effectiveness-enhancing factors from three effectiveness syntheses (Hattie, 2009; Robinson, Hohepa, & Lloyd, 2009; Scheerens, 2016) and a data set of 595 school interventions in Dutch secondary schools (Neeleman, 2019b) shows a meagre overlap between certain policy domains that are present in educational practice - especially in organisational and staff domains - and those interventions currently focussed on in EE research. Vice versa, there are research objects in EE that hardly make it to educational practice, even those with considerable effect sizes, such as self-report grades, formative evaluation, or problem-solving teaching.

How are we to interpret and remedy this incongruity? We know from previous research that educational practice is not always predominantly driven by the need to have an increase in school and student outcomes as measured in cognitive tests (often maths and languages) - the main effect size of most EE. We are also familiar with the much-discussed gap between educational research and educational practice (Broekkamp & Van Hout-Wolters, 2007; Brown & Greany, 2017; Levin, 2004; Vanderlinde & Van Braak, 2009) – two clashing worlds speaking different languages and with only few interpreters around. In this paper, we argue for a number of changes in SI to enhance its potential for improving students' chances in life. These changes in SI refer to the context (2), the classroom and teaching (3), the development of SI capacity (4), the interaction with communities (5), and the transfer of SI research into practice (6).

3.2   Contextually Variable School Improvement

Throughout their development, SI and EE have had very little to say about whether or not 'what works' is different in different educational contexts. This happened in part since the early EE discipline had an avowed 'equity' or 'social justice' commitment. This led to an almost exclusive focus in research in many countries on the schools that disadvantaged students attended, leading to the absence of school contexts of other students being in the sampling frame. At a later time, this situation has changed, with most studies now being based upon more nationally representative samples, and with studies attempting to focus on establishing 'what works' across these broader contexts (Scheerens, 2016).

Looking at EE, we cannot emphasize enough that many findings are based on studies conducted in primary education in English-speaking and highly developed countries - mostly, but not exclusively, in the US (Hattie, 2009). From Scheerens (2016, p. 183), we know that "positive findings are mostly found in studies carried out in the United States." Nevertheless, many of the statistical relationships

established in EE over time between school characteristics and student outcomes are on the low side in most of the meta analyses (e.g. Hattie, 2009; Marzano, 2003) with a low variance in outcomes being explained by the use of single school-level factors or averaged groups of them overall.

Strangely, this has insufficiently led to what one might have expected – the disaggregation of samples into smaller groups of schools in accordance with characteristics of their contexts, like socioeconomic background, ethnic (or immigrant) background, urban or rural status, and region. With disaggregation and analysis by groups of schools within these different contexts, it is possible that there could be better school-outcome relationships than overall exist across all contexts with school effects seen as moderated by school context.

This point is nicely made by May, Huff, and Goldring (2012) in an EE study that failed to establish strong links between principals' behaviours and attributes in terms of relating the time spent by principals on various activities and student achievement over time leading to the authors' conclusion that "…contextual factors not only have strong influences on student achievement but also exert strong influences on what actions principals need to take to successfully improve teaching and learning in their schools" (p. 435).

The authors rightly conclude in a memorable paragraph that,

> …our statistical models are designed to detect only systemic relationships that appear consistently across the full sample of students and schools. […] if the success of a principal requires a unique approach to leadership given a school's specific context, then simple comparisons of time spent on activities will not reveal leadership effects on student performance. (also p. 435)

3.2.1 The Role of Context in EE over the Last Decades

In the United States, there was an historic focus on simple contextual effects. Their early definition thereof as 'group effects' on educational outcomes was supplemented in the 1980s and 1990s by a focus on whether the context of the 'catchment area' of the school influenced the nature of the educational factors that schools used to increase their effectiveness. Hallinger and Murphy's (1986) study of 'effective' schools in California, which pursued policies of active parental disinvolvement to buffer their children from the influences of their disadvantaged parents/caregivers, is just one example of this focus. The same goes for the Louisiana School Effectiveness Study (LSES) of Teddlie and Stringfield (1993). Furthermore, there has also been an emphasis in the UK upon how schools in low SES communities need specific policies, such as the creation of an orderly structured atmosphere in schools, so that learning can take place (see reviews in Muijs, Harris, Chapman, Stoll, & Russ, 2004; Reynolds et al., 2014). Also, in the UK, the 'site' of ineffective schools was the subject of intense speculations for a while within the school improvement community in terms of different, specific interventions that were needed due to their distinctive pathology (Reynolds, 2010; Stoll & Myers, 1998).

However, this flowering of what has been called a 'contingency' perspective did not last very long. The initial *International Handbook of School Effectiveness Research* (Teddlie & Reynolds, 2000) comprises a substantial chapter on 'context specificity' whereas the 2016 version does not (Chapman et al., 2016).

Subsequently, many of the lists that were compiled in the 1990s concerning effective school factors and processes had been produced using research grants from official agencies that were anxious to extract 'what works' from the early international literature on school effectiveness in order to directly influence school practices. In that context, researchers recognised that acknowledging the findings from schools that showed different process factors being effective in different ways in different contextual areas, would not give the funding bodies what they wanted. Many of the lists were designed for practitioners, who might appreciate the universal mechanisms about 'what works.' There was a tendency to report confirmatory findings rather than disconfirmatory ones, which could have been considered 'inconvenient.' The school effectiveness field wanted to show that it had alighted on truths: 'well, it all depends upon context' was not a view that we believed would be respected by policy and practice. The early EE tradition that showed that 'what works' was different in different contexts had largely vanished.

Additional factors reinforced the exclusion of context in the 2000s. First, the desire to ape the methods employed within the much-lauded medical research community – such as experimentation and RCTs – reflected a desire, as in medicine, to be able to intervene in all educational settings with the same, universally applicable methods (as with a universal drug for all illness settings, if one were to exist). The desire to be effective across all school contexts – 'wherever and whenever we choose' (Edmonds, 1979, cited in Slavin, 1996) – was a desire for universal mechanisms. Yet, of course, the medical model of research is in fact designed to generate universally powerful interventions and, at the same time, is committed to context specificity with effective interventions being tailored to the individual patient's context in terms of the kind of drug used (for example one of the forty variants of statin), dosage of the drug, length of usage of the drug, combination of a drug with other drugs, the sequence of usage if combined with other drugs, and patient-dependent variables, like gender, weight, and age. We did not understand this in EE – or perhaps we did comprehend this, but this was not a convenient stance for our future research designs and funding. We picked up on the 'universal' applicability but not on the contextual variations. Perhaps we also did not sufficiently recognise the major methodological issues about randomised controlled trials themselves – particularly the issues that deal with sample atypicality.

Second, the meta-analyses that were undertaken ignored contextual factors in the interests of substantial effect sizes. Indeed, national context and local school SES context were rarely factors used to split the overall sample sizes, and (when they did) were based upon superficial operationalization of context (e.g. Scheerens, 2016).

Third, the rash of internationally based studies that attempted to look for regularities cross-culturally in the characteristics of effective schools, and school systems were also of the 'one right way' variety. The operationalization of what were

usually highly abstract formulations – such as a 'guiding coalition' or group of influential educational persons in a society – was never sufficiently detailed to permit testing of ideas.

Fourth, the run-of-the-mill multilevel, multivariate EE studies analysing whole samples did not disaggregate into SES contexts, urban/rural contexts, or ethnic (or immigrant) background as this would have cut the sample size. Hence, context was something that – as a field – we controlled out in our analyses, not something that we kept in in order to generate more sensitive, multi-layered explanations.

Finally, many of the nationally based educational interventions generated within many Anglo-Saxon societies that were clearly informed by the EE literature involved intervening in disadvantaged, low-SES communities, but with programmes derived from studies that had researched and analysed their data for all contexts, universally. The circle was complete from the 1980s and 1990s research: Specific contexts received programmes generated from universally based research.

It is possible that for understandable reasons, a tradition in educational effectiveness that would have been involved in studying the complex interaction between context and educational processes, and that would have also generated further knowledge about 'what works by context', has eroded. This tradition needs to be rebuilt and placed in many educational contexts and applied in school improvement.

3.2.2 Meaningful Context Variables for SI

What contextual factors might provide a focus for a more 'contingently orientated' SI approach to 'what works' to improve schools? The socio-economic composition of the 'catchment areas' of schools is just one important contextual variable – others are whether schools are urban or rural or 'mixed,' the level of effectiveness of the school, the trajectory of improvement (or decline) in school results over time, and the proportion of students from a different ethnic (or immigrant) background. Various of these areas have been explored – by Hallinger and Murphy (1986), Teddlie and Stringfield (1993), and Muijs et al. (2004) on SES contextual effects, and by Hopkins (2007), for example, in terms of the effects of where an individual school may be within its own performance cycle affecting what needs to be done to improve.

Other contextual factors that may indicate a need for different interventions in what is needed to improve include:

- Whether the school is primary or secondary for the student age groups covered and/or whether the school is of a distinct organizational type (e.g. selective);
- Whether the school is a member of educational improvement networks;
- Whether the school has significant within-school variation in outcomes, such as achievement that may act as a brake upon any improvement journey, or which could, contrastingly, provide a 'benchmarking' opportunity.

- Other possible factors concerning cultural context are:

  - school leadership
  - teacher professionalism/culture
  - complexity of student population (other than SES; regarding inclusive education) and that of parents
  - financial position
  - level of school autonomy and market choice mechanisms
  - position within larger school board/academy and district level "quality" factors

We must conclude by saying that for SI, we simply do not know the power of contextually variable SI.

3.3 School Improvement and Classrooms/Teaching

The importance of the classroom level by comparison with that of the school has so far not been marked by the volume of research that is needed in this area. In all multilevel analyses undertaken, the amount of variance explained by classrooms is much greater than that of the school (see for example Muijs & Reynolds, 2011); yet, it is schools that have generally received more attention from researchers in both SI and EE.

Research into classrooms poses particular problems for researchers. Observation of teachers' teaching is clearly essential to relate to student achievement scores, but in many societies access to classrooms may be difficult. Observation is time-consuming, as it is important (ethically) to involve briefing and debriefing of research (methods) to individual teachers and parents. The number of instruments to measure teaching has been limited, with the early American instruments of the 'process-product' tradition being supplemented by a limited number of instruments from the United Kingdom (e.g. Galton, 1987; Muijs & Reynolds, 2011) and from international surveys (Reynolds, Creemers, Stringfield, Teddlie, & Schaffer, 2002). The insights of PISA studies, and, of course, those of the International Association for the Evaluation of Educational Achievement (IEA), such as TIMMS and PIRLS, say very little about teaching practices because they measure very little about them, with the exception of TALIS.

Instructional improvement at the level of the teacher/teaching is relatively rare, although there have been some 'instructionally based' efforts, like those of Slavin (1996) and some of the experimental studies that were part of the old 'process-product' tradition of teacher effectiveness research in the United States in the 1980s and 1990s.

However, it seems that SI researchers and practitioners are content to pull levers of intervention that operate mostly at the school level, even though EE repeatedly has shown that they will have less effect than classroom or classroom/school-based ones. It should be mentioned that the problems of adopting a school-based rather than a classroom-based approach have been magnified by the use of multilevel

modelling from the 1990s onwards, which only allocates variance 'directly' to different levels rather than looking at the variance explained by the interaction between levels (of school and classroom potentiating each other).

3.3.1 Reasons for Improving Teaching to Foster SI

Research in teaching and the improvement of pedagogy are also needed in order to deal with the further implications of the rapidly growing field of cognitive neuroscience, which has been generated by brain imaging technology, such as Magnetic Resonance Imaging (MRI). Interestingly, the field of cognitive neuroscience has been generated by a methodological advance in just the same way that EE was generated by one, in this latter case, value-added analyses.

Interesting evidence from cognitive neuroscience includes:

- Spaced learning, with suggestions that use of time spaces in lessons, with or without distractor activities, may optimise achievement;
- The importance of working or short-term memory not being overloaded, thereby restricting capacity to transfer newly learned knowledge/skills to long-term memory;
- The evidence that a number of short learning sessions will generate greater acquisition of capacities than more rare, longer sessions -the argument for so-called 'distributed practice';
- The relation between sleep and school performance in adolescents (Boschloo et al., 2013).

So, given the likelihood of the impact of neuroscience being major in the next decade, it is the classroom that needs to be a focus as well as the school 'level'. School improvement, historically, even in its recent manifestation, has been poorly linked – conceptually and practically – with the classroom or 'learning level'.

The great majority of the improvement 'levers' that have been pulled historically are all at the school level, such as through development planning or whole school improvement planning, and although there is a clear intention in most of these initiatives for classroom teaching and student learning to be impacted upon, the links between the school level and the level of the classroom are poorly conceptualised, rarely explicit, and even more rarely practically drawn.

The problems with the, historically, mostly 'school level' orientation of school improvements as judged against the literature are, of course, that:

- Within school variation by department within secondary school and by teacher within primary school is much greater than the variation between schools on their 'mean' levels of achievement and 'value added' effectiveness (Fitz-Gibbon, 1991);
- The effect of the teacher and of the classroom level in those multi-level analyses that have been undertaken, since the introduction of this technique in the mid-1980s, is probably three to four times greater than that of the school level (Muijs & Reynolds, 2011).

A classroom or 'learning level' orientation is likely to be more productive than a 'school level' orientation for achievement gains, for the following reasons:

- The classroom can be explored using the techniques of 'pupil voice' that are now so popular;
- The classroom level is closer to the student level than is the school level, opening up the possibility of generating greater change in outcomes through manipulation of 'proximal variables';
- Whilst not every school is an effective school, every school has within itself some classroom practice that is relatively more effective than its other practice. Many schools will have within themselves classroom practice that is absolutely effective across all schools. With a within school 'learning level' orientation, every school can benefit from its own internal conditions;
- Focussing on classroom may be a way of permitting greater levels of competence to emerge at the school level;
- There are powerful programmes (e.g. Slavin, 1996) that are classroom-based, and powerful approaches, such as peer tutoring and collaborative groupwork;
- There are extensive bodies of knowledge related to the factors that effective teachers use and much of the novel cognitive neuroscience material that is now so popular internationally has direct 'teaching' applications;
- There are techniques, such as lesson study, that can be used to transfer good practice, as outlined historically in *The Teaching Gap* (Stigler & Hiebert, 1999).

3.3.2   Lesson Study and Collaborative Enquiry to Foster SI

Much is made in this latter study of the professional development activities of Japanese teachers, who adopt a 'problem-solving' orientation to their teaching, with the dominant form of in-service training being the lesson study. In lesson study, groups of teachers meet regularly over long periods of time (ranging from several months to a year) to work on the design, implementation, testing, and improvement of one or several 'research lessons'. By all indications, report Stigler and Hiebert (1999),

> lesson study is extremely popular and highly valued by Japanese teachers, especially at the elementary school level. It is the linchpin of the improvement process and the premise behind lesson study is simple: If you want to improve teaching, the most effective place to do so is in the context of a classroom lesson. If you start with lessons, the problem of how to apply research findings in the classroom disappears. The improvements are devised within the classroom in the first place. The challenge now becomes that of identifying the kinds of changes that will improve student learning in the classroom and, once the changes are identified, of sharing this knowledge with other teachers, who face similar problems, or share similar goals in the classroom. (p. 110)

It is the focus on improving instruction within the context of the curriculum, using a methodology of collaborative enquiry into student learning, that provides the usefulness for contemporary school improvement efforts. The broader argument is that

it is this form of professional development, rather than efforts at only *school* improvement, that provides the basis for the problem-solving approach to teaching adopted by Japanese teachers.

3.4 Building School Improvement Capacity

We noted earlier that conventional educational reforms may not have delivered enhanced educational outcomes because they did not affect school capacity to improve, merely assuming that educational professionals were able to surf the range of policy initiatives to good effect. Without the possession of 'capacity,' schools will be unable to sustain continuous improvement efforts that result in improved student achievement. It is therefore critical to be able to define 'capacity' in operational terms. The IQEA school improvement project, for example, demonstrated that without a strong focus on the internal conditions of the school, innovation work quickly becomes marginalised (Hopkins 2001). These 'conditions' have to be worked on at the same time as the curriculum on other priorities the school has set itself and are the internal features of the school, the 'arrangements' that enable it to get its work done (Ainscow et al., 2000). The 'conditions' within the school that have been associated with a capacity for sustained improvement are:

- A commitment to staff development
- Practical efforts to involve staff, students, and the community in school policies and decisions
- 'Transformational' leadership approaches
- Effective co-ordination strategies
- Serious attention to the benefits of enquiry and reflection
- A commitment to collaborative planning activity

The work of Newmann, King, and Young (2000) provided another perspective on conceptualising and building learning capacity. They argue that professional development is more likely to advance achievement for all students in a school, if it addresses not only the learning of individual teachers, but also other dimensions concerned with the organisational capacity of the school. They defined school capacity as the collective competency of the school as an entity to bring about effective change. They suggested that there are four core components of capacity:

- The knowledge, skills, and dispositions of individual staff members;
- A professional learning community – in which staff work collaboratively to set clear goals for student learning, assess how well students are doing, and develop action plans to increase student achievement, whilst being engaged in inquiry and problem-solving;
- Programme coherence – the extent to which the school's programmes for student and staff learning are co-ordinated, focused on clear learning goals and sustained over a period of time;

- Technical resources – high quality curriculum, instructional material, assessment instrument, technology, workspace, etc.

Fullan (2000) notes that this four-part definition of school capacity includes 'human capital' (i.e. the skills of individuals), but he concludes that no amount of professional development of individuals will have an impact, if certain organisational features are not in place. He maintains that there are two key organisational features necessary. The first is 'professional learning communities', which is the 'social capital' aspect of capacity. In other words, the skills of individuals can only be realised, if the relationships within the schools are continually developing. The other component of organisational capacity is programme coherence. Since complex social systems have a tendency to produce overload and fragmentation in a non-linear, evolving fashion, schools are constantly being bombarded with overwhelming and unconnected innovations. In this sense, the most effective schools are not those that take on the most innovations, but those that selectively take on, integrate and co-ordinate innovations into their own focused programmes.

A key element of capacity building is the provision of in-classroom support, or in a Joyce and Showers term, 'peer coaching'. It is the facilitation of peer coaching that enables teachers to extend their repertoire of teaching skills and to transfer them from different classroom settings to others. In particular, peer coaching is helpful when (Joyce, Calhoun, & Hopkins, 2009):

- Curriculum and instruction are the contents of staff development;
- The focus of the staff development represents a new practice for the teacher;
- Workshops are designed to develop understanding and skills;
- School-based groups support each other to attain 'transfer of training'.

3.5   Studying the Interactions Between Schools, Homes, and Communities

Recent years have seen the SI field expand its interests into new areas of practice, although the acknowledgement of the importance of new areas has only to a limited degree been matched by a significant research enterprise to fully understand their possible importance.

Early research traditions established in the field encouraged the study of 'the school' rather than of 'the home' because of the oppositional nature of our education effectiveness community. Since critics of the field had argued that 'schools make no difference', we in EE, by contrast, argued that schools do make a difference and proceeded to study schools exclusively, not communities or families together with schools.

More recently, approaches, which combine school influences and neighbourhood/social factors in combination to maximise influence over educational achievement, have become more prevalent (Chapman et al., 2012). The emphasis is now

upon 'beyond school' rather than merely 'between school' influences. Specifically, there is now:

- A focus upon how schools cast their net wider than just 'school factors' in their search for improvement effects (Neeleman, 2019a), particularly, in recent years, involving a focus upon the importance of outside school factors;
- As EE research has further explored what effective schools do, the 'levers' these schools use have increasingly been shown to involve considerable attention to home and to community influences within the 'effective' schools;
- It seems that, as a totality, schools themselves are focussing more on these extra-school influences, given their clear importance to schools and given schools' own difficulty in further improving the quality of already increasingly 'maxed out' internal school processes and structures; but this might also be largely context-dependent;
- Many of the case studies of successful school educational improvement, school change, and, indeed, many of the core procedures of the models of change employed by the new 'marques' of schools, such as the Academies' Chains in the United Kingdom and Charter Schools in the United States, give an integral position to schools attempting to productively link their homes, their community, and the school;
- It has become clear that variance in outcomes explained by outside school factors is so much greater than the potential effects of even a limited, synergistic combination of school and home influences could be considerable in terms of effects upon school outcomes;
- The variation in the characteristics of the outside world of communities, homes, and caregivers itself is increasing considerably with the rising inequalities of education, income, and health status. It may be that these inequalities are also feeding into the maximisation of community influences upon schools and, therefore, potentially the mission of SI. At least, we should be aware of the growing gap between the haves and the have-nots (or, following David Goodhart, the somewheres and the anywheres) in many Western (European) countries and its possible influence on educational outcomes.

3.6 Delivering School Improvement Is Difficult!

Even accepting that we are clear on the precise 'levers' of school improvement, and we have already seen the complexity of these issues, it may be that the characteristics, attributes, and attitudes of those in schools, who are expected to implement improvement changes, may somewhat complicate matters. The work of Neeleman (2019a), based on a mixed-methods study among Dutch secondary school leaders, suggests a complicated picture:

- School improvement is general in nature rather than being specifically related to the characteristics of schools and classrooms outlined in research;
- School leaders' personal beliefs relate to connecting and collaborating with others, a search for moral purpose and the need to facilitate talent development and generate well-being and safe learning environments. Their core beliefs are about strong, value- driven, holistic, people-centred education, with an emphasis on relationships with students and colleagues. Rather than being motivated by the ambition to improve students' cognitive attainment, which is what school improvement and school improvers emphasize.
- School leaders interpret cognitive student achievement as a set of externally defined accountability standards. As long as these standards are met, they are rather motivated by holistic, development-oriented, student-centred, and non-cognitive ambitions. This is rather striking in light of current debates about the alleged influence of such standardized instruments on school practices, as critics have claimed that these instruments limit and steer practitioners' professional autonomy.
- Instead of concluding that school leaders are not driven by the desire to improve cognitive student achievement as commonly defined in EE research or enacted in standardized accountability frameworks, one could also claim that school leaders define or enact the notion differently. Rather than finding the continuous improvement of cognitive student achievement the holy grail of education, they seem more driven by the goal of offering their students education that prepares them for their future roles in a changing society. This interpretation implies more customized education with a focus on talent development and noncognitive outcomes, such as motivation and ownership. Such objectives, however, are seldom used as outcome measures in EE research or accountability frameworks.
- If evidence plays a role in school leaders' intervention decision-making, it is often used implicitly and conceptually, and it frequently originates from personalized sources. This suggests a rather minimal direct use of evidence in school improvement. The liberal conception of evidence that school leaders demonstrate is striking, all the more so, if one compares this interpretation to common conceptions of evidence in policy and academic discussions about evidence use in education. School leaders tend to assign a greater role to tacit knowledge and intuition in their decision-making than to formal or explicit forms of knowledge

In all, these findings raise questions in light of the ongoing debate about the gap between educational research and practice. If, on the one hand, school leaders are generally only slightly interested in using EE research, this would indicate the failure of past EE efforts. If, on the other hand, school leaders are indeed interested in using more EE evidence in their school improvement efforts, but insufficiently recognize common outcome measures or specific (meta-)evidence on their considered interventions, then we have a different problem. These questions require answers, if we want to bridge the gap between EE and SI and, thereby, strengthen school improvement capacity.

References

Ainscow, M., Farrell, P., & Tweddle, D. (2000). Developing policies for inclusive education: a study of the role of local education authorities. *International Journal of Inclusive Education, 4*(3), 211–229.

Boschloo, A., Krabbendam, L., Dekker, S., Lee, N., de Groot, R., & Jolles, J. (2013). Subjective sleepiness and sleep quality in adolescents are related to objective and subjective measures of school performance. *Frontiers in Psychology, 4*(38), 1–5. https://doi.org/10.3389/fpsyg.2013.00038

Broekkamp, H., & Van Hout-Wolters, B. (2007). The gap between educational research and practice: A literature review, symposium, and questionnaire. *Educational Research and Evaluation, 13*(3), 203–220. https://doi.org/10.1080/13803610701626127

Brown, C., & Greany, T. (2017). The evidence-informed school system in England: Where should school leaders be focusing their efforts? *Leadership and Policy in Schools*. https://doi.org/10.1080/15700763.2016.1270330

Chapman, C., Muijs, D., Reynolds, D., Sammons, P., & Teddlie, C. (2016). *The Routledge international handbook of educational effectiveness and improvement: Research, policy, and practice*. London and New York, NY: Routledge.

Chapman, C., Armstrong, P., Harris, A., Muijs, D., Reynolds, D., & Sammons, P. (Eds.). (2012). *School effectiveness and school improvement research, policy and practice: Challenging the orthodoxy*. New York, NY/London, UK: Routeledge.

Creemers, B., & Kyriakides, L. (2009). Situational effects of the school factors included in the dynamic model of educational effectiveness. *South African Journal of Education, 29*(3), 293–315.

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 37*(1), 15–27.

Fitz-Gibbon, C. T. (1991). Multilevel modelling in an indicator system. In S. Raudenbush & J. D. Willms (Eds.), *Schools, pupils and classrooms: International studies of schooling from a multilevel perspective* (pp. 67–83). London, UK/New York, NY: Academic.

Fullan, M. (2000). The return of large-scale reform. *Journal of Educational Change, 1*(1), 5–27.

Galton, M. (1987). An ORACLE chronicle: A decade of classroom research. *Teaching and Teacher Education, 3*(4), 299–313.

Hallinger, P., & Murphy, J. (1986). The social context of effective schools. *American Journal of Education, 94*(3), 328–355.

Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.

Hopkins, D. (2001). *Improving the quality of education for all*. London: David Fulton Publishers.

Hopkins, D. (2007). *Every school a great school*. Maidenhead: Open University Press.

Joyce, B. R., Calhoun, E. F., & Hopkins, D. (2009). *Models of learning: Tools for teaching* (3rd ed.). Maidenhead: Open University Press.

Levin, B. (2004). Making research matter more. *Education Policy Analysis Archives, 12*(56). Retrieved from http://epaa.asu.edu/epaa/v12n56/

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.

May, H., Huff, J., & Goldring, E. (2012). A longitudinal study of principals' activities and student performance. *School Effectiveness and School Improvement, 23*(4), 415–439.

Muijs, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socio-economically disadvantaged areas: A review of research evidence. *School Effectiveness and School Improvement, 15*(2), 149–175.

Muijs, D., & Reynolds, D. (2011). *Effective teaching: Evidence and practice*. London, UK: Sage.

Neeleman, A. (2019a). *School autonomy in practice. School intervention decision-making by Dutch secondary school leaders*. Maastricht, The Netherlands: Universitaire Pers Maastricht.

Neeleman, A. (2019b). The scope of school autonomy in practice: An empirically based classification of school interventions. *Journal of Educational Change, 20*(1), 3155. https://doi.org/10.1007/s10833-018-9332-5

Newmann, F., King, B., & Young, S. P. (2000). Professional development that addresses school capacity. Paper presented at American Educational Research Association Annual Conference, New Orleans, 28 April.

Reynolds, D., Creemers, B. P. M., Stringfield, S., Teddlie, C., & Schaffer, E. (2002). *World class schools: International perspectives in school effectiveness*. London, UK: Routledge Falmer.

Reynolds, D. (2010). *Failure free education? The past, present and future of school effectiveness and school improvement*. London, UK: Routledge.

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, 2592*, 197–230.

Robinson, V., Hohepa, M., & Lloyd, C. (2009). *School leadership and student outcomes: Identifying what works and why. Best evidence synthesis iteration (BES)*. Wellington, New Zealand: Ministry of Education.

Scheerens, J. (2016). *Educational effectiveness and ineffectiveness. A critical review of the knowledge base*. Dordrecht, The Netherlands: Springer.

Slavin, R. E. (1996). *Education for all*. Lisse, The Netherlands: Swets & Zeitlinger.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press, Simon and Schuster, Inc.

Stoll, L., & Myers, K. (1998). *No quick fixes*. London, UK: Falmer Press.

Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10-year study of school effects*. New York, NY: Teachers College Press.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. London, UK: Falmer Press.

Vanderlinde, R., & van Braak, J. (2009). The gap between educational research and practice: Views of teachers, school leaders, intermediaries and researchers. *British Educational Research Journal, 36*(2), 299–316. https://doi.org/10.1080/01411920902919257