Chapter 1

School Development Approaches Over Time: Strengths, Limitations, and the Need for a New Approach

Rose M. Ylimaki and Lynnette A. Brunderman

Abstract Across the globe, we observe policy trends towards evidence-based school development, “scientific” research, and increasingly centralized curriculum, all occurring amidst growing digitalization and demographic changes resulting in increasingly pluralistic schools and communities. As a result of these policy pressures, many universities and other educational organizations have proposed various evidence-based school development models or projects aimed at continuous improvement. In this chapter, we contest evidence produced from quasi-experimental research designs with other empirically tested evidence. We critique several established school development programs across the U.S. Finally, we propose that in order to have school capacity for continuous development, we need a school development process that is contextually-based and able to consider, reflect upon and use data from evidence-based programs and other sources as appropriate for school-identified problems, and that schools must be able to evaluate their own programs and processes in their particular contexts, with particular populations. For this, we need school development grounded in understandings of education and leadership as well as evidence. For our purposes, we define school development as a continuous growth process for school teams supporting education amidst tensions between policy expectations for use of research-driven evidence and the needs of students in increasingly diverse contexts.

Keywords Global policy trends · Cultural diversity · Democratic and humanistic values · School development

R. M. Ylimaki (✉)
Northern Arizona University, Flagstaff, AZ, USA
e-mail: rose.ylimaki@nau.edu

L. A. Brunderman
University of Arizona, Tucson, AZ, USA
e-mail: lbrunder@arizona.com

© The Author(s) 2022

R. M. Ylimaki, L. A. Brunderman (eds.), *Evidence-Based School Development in Changing Demographic Contexts*, Studies in Educational Leadership 24, https://doi.org/10.1007/978-3-030-76837-9_1

The book is about leadership capacity for continuous school development and improved student outcomes for culturally diverse youth in traditionally underserved communities amidst the contemporary evidence-based policy context. Across the globe, we can observe policy trends toward evidence-based school development, instrumentalism, evaluation, “scientific” research, and increasingly national or centralized curriculum, all of which is occurring amidst growing digitalization with a knowledge economy as well as global and internal demographic changes toward increasingly pluralistic schools and communities. As a result of policy pressures on schools and related pressures on higher education to demonstrate impact of research, many universities and other educational organizations across the U.S. and elsewhere have proposed various evidence-based school development models or projects aimed at continuous improvement. While many of these projects or models have been empirically tested and are working toward a so-called gold standard of strong evidence, few have strong evidence from experimental design promoted in What Works Clearinghouse and a number of research grants. Approaches grounded in standards, assessments and other central regulations seek to manage school practice, but Bryk et al. (2015) and others argue that these solutions are often oversold, under-implemented, and soon replaced with little learned in the process. Closely related, other evidence-based practices for school development involve creating potentially effective and replicable programs, evaluating them (ideally according to What Works Clearinghouse in randomized controlled experiments) and providing incentives for schools to use those that are found to be effective. Rather, Bryk et al. (2015, p. 468–469) notes that while clinical trials are worth doing in education, they are a very slow and expensive process, and such studies are not likely to be the primary resource for improving schools anytime soon. Bryk et al. go on to argue that randomized trials may just report effect sizes, indicating that a program can work, but it may not show how to make it work reliably over diverse contexts and populations. Moreover, from our perspective, clinical trials do not take into account leadership capacity, teachers’ pedagogical capacity, or increasingly diverse student needs and interests. In order to have school capacity for continuous development, we need a school development process that is contextually-based and able to consider, reflect upon, and use data from evidence based programs and other sources as appropriate for particular problems and conduct their own experiments or assessments about how programs and processes work in particular contexts and with particular populations. For this, we need school development grounded in understandings of education and leadership as well as evidence.

For present purposes, we define school development as a continuous growth process for school teams supporting education amidst tensions between policy document expectations for use of evidence from experimentally designed programs and the needs of students in increasingly diverse school contexts. This volume features an ongoing project developed for culturally diverse Arizona schools and most recently applied to the South Carolina context as well as similar school development projects in Australia, Germany, and Sweden. School development projects are considered in relation to current national and cross-national policy trends toward evidence-based policies and changing demographics and in relation to particular contexts. We conclude with our thoughts about the need for further work in the area of school development.

Evidence-Based Policy Trends

Historically in the U.S., policy documents under both Republican and Democratic administrations have supported the need for evidence from externalized evaluations (standardized assessment data) to guide school decisions. Additionally, since 2002, there has been a drastic rise in the call for education research funded by the Office of Educational Research and Improvement (OERI) that addresses causal questions using random assignment designs [prior funding for randomized controlled trials represented 5% of federal funding for education research compared to 75% by 2002 (Morrison, 2012)]. Currently, funding applications for U.S. Department of Education grants must include research designs that are based upon prior studies with “strong evidence” (explicitly defined by large-scale quantitative studies with randomized controlled trials or quasi-experimental designs). In the international arena, multinational organizations (the Organization for Economic Co-operation and Development (OECD) and the World Bank) have also made evidence-based policymaking a priority, both in their own work as well as for their members (Wiseman, 2010).

In recent years, we can observe global changes in educational policies and governance systems with increased curriculum centralization, the advent of externalized evaluation policies and the increasing scrutiny of educational organizations at all levels, particularly public schools. More specifically, many nation states, including the U.S., have experienced both increased centralization (e.g., curriculum, evaluations) as well as increased decentralization (school-based decisions about improvements). New bureaucracies have emerged in relation to neoliberal policies, all of which demand strong leaders who can mediate among many different (and sometimes conflicting) policies, diverse student needs, and democratic education values. Further, school leaders now must work collaboratively within and between governance systems or levels, balancing evidence-based values and the humanistic values of education for increasingly diverse students. In Fig. 1.1 that follows, we illustrate some of the historical policy changes that have affected schools and leaders in the U.S. and internationally.

As the timeline demonstrates, evidence-based policies are not entirely new, however the political pressures have created more stringent external oversight. We can observe that “new” evidence-based policies can be categorized in terms of three interrelated intents, (1) to measure school quality on standardized tests, (2) to create equality among schools, and/or (3) to increase state control of schooling (Wiseman, 2010). For a full discussion of recent evidence-based policy intentions, see Wiseman (2010).

The most popular reason for using evidence as a basis for policymaking is that evidence provides an indicator of quality in terms of how much someone has learned or how much impact a certain educational technique has on students (Wiseman, 2010). Here the underlying assumption is that the more students learn, the more they know. The more students know, the better their test performance (a key source of evidence) will be. The better the students’ test performance is, the better the teacher or school is. Most often summative assessments in the form of

Fig. 1.1 Timeline of evolution of education trends and policies

norm-referenced standardized tests are used to summarize how well an individual or an organization is performing. This logic is flawed because these standardized assessments are normed on a bell curve designed to sort individuals; students with scores reflecting a single test question difference can be designated as deficient. Standardized tests are not designed to measure a student's mastery of content (e.g., Bond, 1995; Decker, 2008; Reynolds et al., 2010). Yet, summative assessments are a primary tool to measure educational quality and represent the final sum of knowledge that a student has been taught (or has learned) up to a particular point. It is important to note that with summative assessments, there is no opportunity for feedback, reflection, planning or further instruction to students. Moreover, norm-referenced standardized assessments have long been criticized for privileging children from White middle class households and communities (Reynolds et al., 2010). Yet the evidence resulting from their state, national or international test performance informs policymakers about the perceived quality of education.

In many nations, including the United States, there are a variety of external evaluations (state tests) given in different states and sometimes in different districts. For example, in Arizona, students currently take the AZMerit test of knowledge and skills. There is no longer a graduation exam¹ in Arizona, but some school districts additionally require all students to take the ACT exam. In South Carolina, beginning in grade 3, students take the South Carolina Palmetto Assessment of State Standards (SCPASS) in English/Language Arts, Mathematics, Science and Social Studies. The High School Assessment Program (HSAP) is first administered in grade 10, and South Carolina high school students must pass this test in order to receive a high school diploma. National tests, such as the National Assessment of Educational Progress (NAEP), provide nationally representative achievement information in core subjects. International assessments that are widely taken in many countries (e.g. PISA, TIMSS) have often driven national policy decisions toward an increased use of tests as evidence of school quality or success. Although there is a range in the extent to which the results of these assessments directly affect individual students, the average student scores for schools, districts, states and nations as a whole are widely accepted indicators of educational quality and have a major impact on educational policy-making (LeTendre et al., 2001).

The degree of equality attained in an educational system is another reason frequently cited for assessing the performance of students and schools. By seeing who performs highly and who does not, it is possible for teachers, administrators and policymakers to determine disparities among individual students, classrooms, and schools (Heilig and Darling-Hammond, 2008; Darling-Hammond, 2010). The logic behind using evidence to demonstrate equality suggests that policymakers can design policies and laws to close achievement gaps. For example, Arizona, like many other states, has a wide achievement gap between students from various social and economic groups. The state test (currently AZMerit) shows that this gap exists

¹Arizona is in the process of changing vendors for a state test again. In the last 10 years, this will be the third test change.

and the extent to which it exists in particular content areas. Arizona politicians have used this evidence to reward high performing schools and to punish low performing schools. Analysis of standardized achievement scores clearly falls along socioeconomic lines across the nation (Perry et al., 2010; Baker and Johnston, 2010; Cunningham and Sanzo, 2002) and has contributed to the rise of charter schools, vouchers and public mechanisms to ensure greater equality among schools. Thus, politicians—with evidence from state tests— have passed legislation and policy targeting underperforming schools, requiring shifts in the curricular and classroom focus in order to close those achievement gaps. Under NCLB, if disaggregated data for each student subgroup (i.e., race, gender, language proficiency, etc.) do not demonstrate adequate yearly progress, then the U.S. Department of Education may withhold federal funding from the state or district within the state (McDermott and Jensen, 2005). In 2015, the reauthorization of NCLB resulted in the Every Student Succeeds Act (ESSA), which returned some control to states to determine their goals. Students are still required to be tested annually in grades 3 through high school, but the federal government removed many of the prescriptive pieces of the legislation (U.S. Department of Education, 2016). However, Arizona (as in many states) has not changed state legislation to reflect the greater flexibility of ESSA. It should also be noted that in Arizona, the state test was designed to have both criterion-references components and others that served as norm-referenced items by testing standards at, above, and below the grade level for which the test was designed.

Historically, federal and many state policy makers have attempted to address inequities through legislation. For example, the Coleman Report from the 1960s suggested that family and peer influences, not school resources, are the most important determinants of student performance (Baker and LeTendre, 2005). The Coleman Report was conducted in the wake of the court case *Brown vs. Board of Education of Topeka* that ruled that racial segregation of children in U.S. public schools was unconstitutional. “Evidence” from the landmark court decision and the latter Coleman Report changed the course of equity-driven policy-making in the United States and in many other countries. The Coleman Report was part of President Johnson’s strategy in the Great Society program to level the playing field by identifying evidence of inequality in schooling. Coleman surveyed over 560,000 students across the U.S., a methodology partly due to the fact that the system was highly decentralized. And while recent externalized evaluation and curriculum policies have shifted toward centralization (e.g. Common Core curriculum), administration of education in the U.S. remains state-based. During the time of the Coleman Report, Johnson and other policymakers thought they could address inequality with additional resources for low-socioeconomic schools as a strategy for improvement. However, while they found unequal resources available according to class and race, Coleman’s team found evidence that some equity indicators like resources did not predict achievement. Variations in family background and outside-of-school environment affected achievement at least as much as variations in school resources and quality. In other words, some policy impacts cannot be controlled by policymakers.

Closely related, Mexican American parents successfully sued school districts in California (*Roberto Alvarez vs. the Board of Trustees of the Lemon Grove School District*, 1931, and *Mendez vs. Westminster*, 1946), both of which attempted to segregate Spanish-speaking Mexican American students in inferior “Americanization” schools. The 1968 passage of the Bilingual Education Act (Title VII of the Elementary and Secondary Education Act) provided federal protections for culturally and linguistically diverse students. The Bilingual Education Act built from the *Brown v. Board of Education* case and the civil rights movement and promoted instruction that was both in a student’s native language and culturally responsive. The act was far reaching, but it (and related state policies) remain at tension with the need to educate students in a common language and other common understandings and civic dispositions that create a nation state. For example, today’s school leadership teams must balance tensions between recent shifts toward curriculum centralization in the Common Core (and state versions thereof) and related externalized evaluation policies designed to measure student performance.

Policy makers control school funding; at state and national levels, test results are increasingly tied to funding, with the most recent policy wave beginning with the No Child Left Behind Act and its requirement for schools to make adequate yearly progress. Because the U.S. Department of Education has the responsibility for evaluating schools’ and states’ performance, it can and does use test performance as evidence and a means to control the curriculum schools adopt, the content teachers teach and other components of schooling (Wiseman, 2010). Since the No Child Left Behind Act, schools identified as not making Adequate Yearly Progress (AYP) over a series of years have been in danger of losing federal funding, and as such this has become vital policymaking evidence for school systems in every state. Thus, some would argue that policymakers have opted toward evidence-based approaches over which they have control.

Globally, we can observe a policy convergence toward such compliance requirements, curriculum centralization and rankings for school quality and performance (Pilton, 2009; Steiner-Khamsi, 2006). As Pilton (2009) explained, there are several ways that policy convergence contributes to evidence-based policies, curriculum centralization, and comparisons or rankings of schools, states, and nations. One way is through imposition. Imposition is typically a result of political demand or pressure (e.g. Organization for Economic Cooperation and Development/OECD agenda, World Bank) and requires compliance with international policy or institutional arrangements. There is no legal obligation to participate in the Programme for International Student Assessment (PISA) with OECD; however, at OECD urging, member nations comply with evidence-based policies. Further, as Meyer (1977) and other institutional theorists would argue, when OECD identifies nation states with “high performing” schools, others emulate these practices (e.g. centralized curriculum) as they are legitimized. Such evidence-based policy convergence also contributed to research trends in funding applications, particularly research methods that are considered “scientific” and measure what works to improve schools.

Research Trends

In the U.S., federal grants legitimize “scientific” research designs and methods similar to the medical field that measure various interventions aimed at improved student outcomes on standardized tests, including randomized controlled trials and quasi-experimental designs. Since the No Child Left Behind Act in 2002, education research funded by the Office of Educational Research and Improvement (OERI) that addresses causal questions using random assignment designs increased from 5% to 75% (Morrison, 2012). In current funding applications for the U.S. Department of Education grants, researchers must demonstrate that their research designs are based upon prior studies with “strong evidence” explicitly defined by large-scale quantitative studies with randomized controlled trials or quasi-experimental designs that primarily measure what works in terms of gains in student outcomes. That is, federally funded research channels future research in a particular and similar direction, and this research is considered legitimate with “strong evidence”.

A number of influential scholars have argued persuasively for the use of such evidence to inform educational practice. Slavin is one of the most frequently cited proponents for the use of research and practice similar to the medical field. Using his Success for All project as an example, Slavin (2010) argues for the importance of studies that seek to make causal conclusions that include correlational and descriptive dimensions as he used in Success for All, one of the innovations featured in What Works Clearinghouse with strong evidence of effectiveness. The What Works Clearinghouse is sponsored by the Institute of Educational Science to provide educators with interventions designed and tested with scientific research that demonstrates effectiveness to improve student outcomes. The push for evidence-based policies is evident around the world. As noted earlier, this trend of convergence toward evidence-based policies has particular contextual challenges in nation states with federal systems of education and in schools serving increasingly diverse populations of students.

Australia, Germany, and the United States have a history of a state-based system with regards to education, but all three countries are experiencing an increasing centralization of curriculum, accountability systems, and evidence-based policy trends. Sweden has a much longer history of a centralized national curriculum; however, recent national policy has shifted to focus on equity within and between schools and municipalities. All of these policy trends have developed amidst changing demographics. Australia, for example, has experienced changing migration patterns from Europe to Asia and Africa. Since the 1970s, Sweden, too, has experienced significant demographic changes due to immigration. Demographic changes in the United States are due to both population migrations and refugees, as well as internal demographic changes. In certain ways, globalizing trends toward centralized curriculum and accountability policies are at tension with the needs of increasingly diverse students in schools.

Education and Humanistic Values

While diversity has increased and changed since Dewey wrote *Democracy and Education* with his philosophy of education, we see parallels with the contemporary situation for continuous growth in school development. Over the course of his career, Dewey (1897a, b, 1916, 1938) encouraged continuous growth and lifelong learning. Dewey (1916) argued, in particular, that an education which only emphasizes the achievement of “external aims” hinders students’ capacity for continuous growth and leads students to view schooling as something to end as quickly as possible. In terms of school development, this might suggest that reliance on evidence from standardized test scores, grades, school letter grades, etc. inhibit school teams’ ability to evolve and grow, resulting in seeing school development as a destination rather than a journey. In contrast, in a Deweyan notion of education and democratic growth, experience is key and growth is a continuous lifelong venture. Additionally, for Dewey, reality is constantly changing and no one has a monopoly on knowledge; this means that democracy requires everyone to continuously grow and adapt to changing conditions.

Dewey (1897/2000) sought to conceptualize education in an earlier time of social and political change. More specifically, Dewey (1900) discusses unifying the student with other students so that the school “gets a chance to be a miniature [democratic] community, an embryonic society. This is the fundamental fact, and from this arise continuous orderly sources of instruction” (p. 32). In our project, we incorporated these education theory perspectives and educative values for building leadership capacity in school development to include thinking, growth, reflection, and pedagogical interaction, as well as an emphasis on lifelong learning and growth. It is important to note that, like Dewey, we did not see the values as fixed but rather as guides to build leadership capacity.

It should also be noted that Dewey (1916) understood the significance, nature and utility of subject areas, or outcomes. In fact, in *Democracy and Education* (1916), Dewey addresses subjects such as geography, history, physical and social studies, as well as “play and work” and vocational aspects of education. He reinforces the aims of education for social direction, preparation, formation, and reconstruction. Thus, education is “based upon a consideration of what is already going on; upon the resources and difficulties of existing conditions” (p. 110), all of which require reflection and development or growth (democracy). More recently, Biesta (2010) renewed attention to democracy, critiquing contemporary “evidence-based” reforms and emphasizing how use of certain evidence-based programs threatens to replace professional judgment and the wider democratic deliberation about the purpose, outcomes and pedagogy of education. Biesta argues for a value-based education rather than an evidence-based education. “Calling the idea of value-based education an alternative is not meant to suggest that evidence plays no role at all in value-based education but is to highlight that its role is subordinate to the values that constitute practices as educational practices” (Biesta, 2010, p. 493). This project recognizes that education values, concepts and aims must interface with values and aims of educational organizations within increasingly diverse communities.

Changing Demographics

Currently, White people constitute the majority of the U.S. population (62%); however, the percentage is expected to fall below 50% by 2060 with Hispanic populations to experience the largest increase (Colby and Ortman, 2015). In Arizona, where we initially designed the school development project, Latino/Latina made up 31% of the population, but the traditionally marginalized non-White populations grew by 62.4% between 2000 and 2015 (Stepler and López, 2016). Poverty levels in Arizona averaged 16.4% with a range from 13.7% in Santa Cruz County to 36.2% in Apache County (Datausa, 2018).

In 2018, we extended the Arizona project to South Carolina, a southeastern U.S. state that has historically reflected a diverse population. At one time, South Carolina was recognized as a super-minority African-American state; current demographics continue to show increasing diversity with a rapidly growing Hispanic and refugee population (Colby and Ortman, 2015). In fact, South Carolina has seen the most rapid growth in Hispanic population of any state in the period between 2000 and 2010 (U.S. Census Bureau, 2011). During this 10-year period, the state saw a 147.9% increase in its Hispanic population (U.S. Census Bureau, 2011). Between 2000 and 2014 the Hispanic Population grew by 172% (Stepler and López, 2016). In addition to the shifting demographics, poverty levels remain high in South Carolina where 17.2% of children in the state live in poverty compared to 15.1% nationally (U.S. Census Bureau, 2018a, b). Minority and low-income students have historically been at-risk for school failure (Reardon et al., 2008). Student achievement data in South Carolina, including those high-needs schools along the I-95 corridor, corroborates the relationship between student demographics and student academic performance (South Carolina Department of Education, 2018).

Cultural diversity continues to bring multiple experiences, values, and views of knowledge to educational institutions and the academics within them. Since the 1960s there has been a significant increase in the number and diversity of immigrants coming to the U.S. The U.S. Census Bureau notes that the increasing number of international migrants will make the country “a ‘plurality’ of racial and ethnic groups” (Colby and Ortman, 2015, p. 9). For example, the Hispanic population has grown from 6.5% of the U.S. population in 1980 to 17.6% in 2015 (Pew Research Center, 2015) while the Two or More Races population is projected to grow by 225% between 2014 and 2060 (Colby and Ortman, 2015). In the next section we consider the strengths and limitations of mainstream school development projects.

Strengths and Limitations of School Development: Research and Popular Models

Over the past five decades, school effectiveness and school improvement research has grown substantially around the globe. Many scholars, often in partnership with local school districts or municipalities, have designed and evaluated the

effectiveness of various approaches to school development. Across the literature, we see an evolving approach to school development, beginning with an organizational health and development approach to studies of an application of findings from effective schools literature conducted in the wake of the Coleman Report, and a focus on change and comprehensive school reform to systems reform.

Early school improvement studies were most often aimed at *organizational* development, grounded in organizational theory (e.g. Lewin, 1947) as well as studies of effective schools (e.g. Edmonds, 1979). Lewin and others (e.g. McGregor, 1960; Miles, 1967) developed experiments to test the influence of the organization on group dynamics, culture and productivity. Miles (1967) applied concepts of organizational health to schools, looking at the relations between the organizational condition of schools and the quality of education they provide. His ten dimensions of organizational health included goals, the transmission of information, and the ways in which decisions are made, the effective use of resources, cohesiveness and morale, ability to deal with growth and change- notions of innovativeness, autonomy, adaptation to the environment and problem-solving. Early organizational development efforts were primarily focused on behavior changes aimed at efficiency with later studies to include an emphasis on how to “humanize” the organizational context. Here we see an emphasis on large-scale national reforms and school leader development (e.g. McLoughlin, 1990). Interestingly, scholars often provided evidence or feedback to school leaders and teachers in the form of survey results. Moreover, scholars who applied organizational development to education identified problems with top-down, externally developed projects that did not take into account the school context but contributed to understanding the leader’s contribution to curriculum, instruction, and organizational change.

Although the U.S. Supreme Court’s decision in *Brown vs. Board* did not achieve school desegregation on its own, the ruling (and steadfast resistance to it across the South and elsewhere) fueled the civil rights movement in the United States. As a prominent example, in 1955, a year after *Brown vs. Board of Education*, Rosa Parks refused to give up her seat on a Montgomery, Alabama bus. Her arrest sparked the Montgomery bus boycott that would lead to other boycotts, sit-ins and demonstrations, many of which were led by Martin Luther King, Jr.

We see evidence-based perspectives with less explicit attention to race and civil rights prominently displayed in later studies from the 1990s but these studies were influenced by the decentralization of schools and the increased prominence of site-based management in the U.S. and elsewhere. School development projects were frequently practitioner led during the 1990s in the UK and the US. During the 1980s, school-based review or self-evaluation became a major strategy for managing change and school development. The empirical support for school-based school reforms was mixed. For these reasons, school improvement during this time period was often seen as implementing an intervention or engaging in action research projects. In many countries, including the United States and Australia, it was also driven by federal funding to address passage of Title I legislation and Australia’s mandate for school improvement councils. During this timeframe, in the United States, the publication of *A Nation at Risk* (1983) initiated what has become a widespread discourse about schools in need of improvements. For example, with this report, we

see the earliest indication of the contemporary pressure toward externalized accountability and change mandating school leaders to understand and apply concrete strategies for change for which they are held accountable.

Thus, in the early 1990s, in the U.S. and elsewhere, school development literature focused on concrete strategies for the management and implementation of change at the school level. The emphasis on change emerged from previous research that indicated the need for more concrete approaches to school improvement as well as the increased need to prepare children for a world engaged in complex and rapid change. While earlier historical times also witnessed educational concerns about schools and education, this particular set of policies and resulting studies of school improvement focused attention on internal school practices in relation to the policy pressures. National policies (Goals, 2000) promoted the idea that school site-based management would free schools from the confines of external (district) control over substantive decisions and pedagogy. At the same time, national and state governments assumed more control over school improvement efforts, a control that increased with the advent of comprehensive school reforms.

Some of these comprehensive school reforms included Levin's Accelerated Schools (1987, 1998) Comer's School Development Program (1996), and Slavin's Success for All (Slavin et al., 1996), mentioned earlier. These whole school design approaches combined elements from the school effectiveness and school improvement research and most often focused in varying degrees on school structures, interpersonal communications, professional development, explicit use of diverse measures of success, and curricula. Some whole school reform models, including Reading Recovery (Clay, 1993) and Success for All were approved for use in What Works Clearinghouse with versions of the models used in many different countries. This meant that schools could receive additional federal funding if they applied one of these particular reform models. Evidence to date indicates a mixed degree of effectiveness from these large scale, externally developed programs (Borman et al., 2003); however, others concluded that locally developed programs were even less likely to result in achieving initially desired outcomes (Nunnery, 1998). In particular, success seemed to be much more elusive in large urban districts. Thus, and in response to the growing pressures for accountability from externalized evaluations and in response to international institutions and comparisons, school development approaches shifted toward system level change and building leadership capacity.

School development aimed at systems change and building leadership capacity for these changes focuses on collaboration and networking across schools and districts. These arguments were developed based upon research in the U.S., the United Kingdom (UK), and elsewhere. According to Harris and Chrispeels (2008) and others, district reform and network building, including professional learning communities, need to occur simultaneously. The essential linkage is provided by an emphasis on building leadership capacity. Examples of such efforts include the National College of School Leadership in the UK (NCSL) and the Carnegie Foundation Project ILead in the U.S. In both of these approaches, the emphasis is on networking complemented by an increasing emphasis on leadership. Closely related, there is a growing body of research on school development and change, including Elmore's

(2004) study of successful school districts in California. In Elmore's (2004) study, participating schools showed a much greater clarity of purpose, a much greater willingness to exercise tighter controls over decisions about what would be taught and what would be monitored as evidence of performance, and a greater looseness and delegation to the school level of specific decisions about how to carry out an instructional program.

Stringfield and Yakimowski-Sreblick (2005) reported on district reforms in historically very low-performing, 90% minority Baltimore City Public Schools. Results from the Baltimore reform situation were mixed with early gains in student outcomes across most schools. However, as accountability systems and policies changed, the district and many of its schools situated in diverse, traditionally marginalized communities experienced many challenges. Early accountability systems did not produce measurable gains. Although some gains were made (graduation rates, outcomes), schools that had been reconstituted and turned over to for-profit management systems showed little to no gains, resulting in a need to be reconstituted again. Teachers and principals required additional professional development to gain skills in interpreting complex data. The study did not indicate any efforts to support teachers in terms of democratic values and practices or culturally responsive pedagogy in culturally diverse settings. Additionally, such reform efforts have not focused exclusively on the role of districts and local authorities in school development but also included other partners or entities, including state and/or national organizations and for-profit partners.

Hopkins (2011) summarized recent key variables in regional systems approaches to school improvement networks that contribute to student achievement. The variables were: (a) a clear and comprehensive model of reform; (b) strong leadership at the regional level; (c) substantive training related to the goals of the programme; (d) implementation support at the school level; and (e) an increasingly differentiated approach to school improvement. Across these studies and projects, the primary aim has been to link school improvement to student learning outcomes. We also see a strong focus on leadership with approaches focused on transformational leadership and change as well as instructional leadership (Leithwood et al., 2004) with more explicit attention to relations among multiple levels of leadership. Scholars are now working to understand the dynamics of improvement working simultaneously within and between levels (Barber, 2009).

Global Interest in School Improvement

Interest in school improvement knowledge is global, including understanding about how to mediate between international benchmarking and other evaluation policies and how to decide among strategies appropriate for both school and system reform (Hopkins, 2013). Here it is important to note that school development is a multi-level process, including the relations among individual schools and districts or municipal systems and state or nation-level systems (Datnow, 2006). Further,

systems are culturally and historically situated; diverse nation states, nations and communities have, over time, developed very different systems for providing education to their children. Further, the systems and levers for systems improvement will vary greatly by national and local context. For example, in Hong Kong (Johnson, 1997), the state determines the core curriculum and the funding level per student. Under the state are a diverse series of school governing bodies, including churches, workers' unions, and so forth. The complexities of change in such a system are dramatically different from those of the United States. As noted in later chapters, schools and districts/municipalities vary considerably according to regional and state context as well as national context. At the same time, we can see growing similarities between school development efforts in Sweden, Germany, Australia and elsewhere. School development projects in these countries will be described later in this book.

Recent studies have examined highly effective educational systems across national contexts. According to Fullan (2009), Hargreaves (1982) and others, highly effective educational systems develop and disseminate clarity on goals and on standards of professional practice, ensuring that student achievement is the central focus of systems' schools; as a consequence, they (1) locate the enhancement of the quality of teaching and learning as central themes in systems improvement, (2) partially achieve their success through selection policies that ensure only highly qualified people become teachers and educational leaders, (3) put in place ongoing and sustained professional learning opportunities that develop a common practice, (4) emphasize school leadership and high expectations that focus on learning, (5) have procedures to enable this and provide timely data for feedback, (6) intervene early at the classroom level to ensure student performance, (7) address inequities in student performance through good early education and direct support for students, and (8) establish system-level structures that link together the various levels of the system and promote disciplined innovation as a consequence of thoughtful professional application of research. While these characteristics are helpful, they do not illustrate the complexity of systems change at multiple levels and in different contexts or the shift from implementation of standardized programs toward a new professionalism among leaders at different levels. Understandings of systemic school development continues to evolve and develop globally along with research understandings of how to scale reform.

Recent U.S. Leadership and School Improvement Programs

More current formal leadership development innovations or programs aimed at improvements in persistently underperforming schools promote such an evidence-based approach to effective leadership and school development, including most prominently the Virginia School Turnaround Specialist Program (STSP), the Mass Insight and Research Institute model based in New York, and the Chicago Reconstitution Effort. For example, the Virginia STSP is an intervention for

principals with a focus on effective leadership practices with effective practices identified as mediators between leadership practice and gains in student outcomes, including data literacy, professional learning, motivation, and curriculum mapping as well as use of evidence-based strategies from the business field, including the development of a 90-day plan for rapid improvement, implementation support, long-term strategic planning and on-site visits. Since 2004, the Virginia STSP program has provided 95 principals with training in business strategies as well as individual coaching to school leaders in more than 82 school districts in numerous states, including Pennsylvania, Illinois, Florida, Missouri, Louisiana, Arizona, New Mexico, Nevada, Utah, Colorado, Texas, Ohio, and the Dakotas, as well as Virginia (Darden/Curry Partnership for Leaders in Education, *ret.* 2019). According to the Virginia STSP report (Harriman, *ret.* 2019), 46% of participants (44) made AYP compared to only 16% (15) prior to participation in the project.

The Mass Insight and Research Institute's project for rapid school improvement (Calkins et al., 2007) proposes a similar evidence-based focus on improvement but aligns schools and service providers into clusters of three to five low-performing schools. Districts and states commit to flexible operating conditions for zone schools with an emphasis on people (recruitment and retention), extended time, money or budget allocation, and program implementation of a rigorous standards-based curriculum and effective leadership practices (e.g. culture building, data literacy, professional learning communities). Results from the Partnership Zones indicated that two-thirds of participants reported gains and one-third reported declines in school performance. Researchers in the School Turnaround Group, a division of Mass Insight Education (2012), attributed the declines in performance to loose coupling between schools and districts.

Chicago's Academy for Urban School Leadership (AUS) drew on the Mass Insight project to improve student achievement in participating schools, including attention to positive school culture; parent engagement; setting goals; shared responsibility for achievement; standards-based, college-prep K–12 curriculum; aligned assessment systems; and engaging personalized instruction. Results indicated that some schools have attained high performance on district benchmarks; however, there were concerns about the sample of students included in the testing. Hood and Ahmed-Ullah reported that these schools have “pushed out the lowest performing children who could not attain the benchmark scores, thus artificially elevating their scores” (Hood and Ahmed-Ullah, 2012, p. 1). Surprisingly, despite the emphasis on instructional leadership in school effectiveness studies, education theory (Dewey, 1916, 1938; Slavin, 2008) with an emphasis on democratic growth and pedagogy has received little attention in educational leadership studies.

While not explicit, all of these innovations imply a primarily closed system for implementation of the innovation or program, meaning that if school leaders apply understandings from effective leadership and school development research within schools and decrease the use of other practices, schools will improve as measured on state tests. In these descriptors, we also see a void in the humanistic values of education and an ontology of education emanating from the later enlightenment and romantic heritage that impacted Dewey's early work (pp. 29–35). In other words,

we see traditional education values and pedagogical methods that support being and becoming in a democratic way of living being replaced by numerical evidence-based values and data analysis methods that support the use of externally developed programs proven to improve student outcomes on standardized tests. Evidence-based programs value replication while Dewey values choice based on context and perspectives of the collective. Dewey values the process of growth (being and becoming) while evidence-based traditions value the outcomes. We return to this point in Chap. 6.

Moreover, despite recent demographic shifts, cultural relevance has not been explicitly addressed in the intervention models reviewed above. In other words, none of these popular evidence-based interventions explicitly considered the humanistic values of education for continuous growth and democracy. The project featured in this volume was designed to balance evidence-based values with the humanistic values of education for increasingly culturally diverse children in schools across the globe (Ylimaki et al., 2019).

In the remainder of this volume, we further highlight evidence-based policies in relation to democratic values in culturally diverse settings. Our discussion of evidence-based policy trends is culturally and historically situated. We then present lessons from school development projects, particularly including AZiLDR and lessons learned from that project as well as similar approaches across the globe. The volume concludes with implications for building capacity among scholars/researchers and practitioners in increasingly pluralistic and democratic communities with attention to leadership preparation, practice and policy.

References

- Baker, M., & Johnston, P. (2010). The impact of socioeconomic status on high stakes testing reexamined. *Journal of Instructional Psychology*, 37(3), 193–199.
- Baker, D., & LeTendre, G. K. (2005). *National differences, global similarities: World culture and the future of schooling*. Stanford University Press.
- Barber, M. (2009). From system effectiveness to system improvement: Reform paradigms and relationships. In *Change wars* (pp. 71–94). Solution Tree.
- Biesta, G. J. (2010). Why ‘what works’ still won’t work: From evidence-based education to value-based education. *Studies in Philosophy and Education*, 29(5), 491–503.
- Bond, L. A. (1995). *Norm-referenced testing and criterion-referenced testing: The differences in purpose, content, and interpretation of results*. ERIC.
- Borman, G. D., Hewes, G., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125–230.
- Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). *Learning to improve: How America’s schools can get better at getting better*. Harvard Education Press.
- Calkins, A., Guenther, W., Belfore, G., & Lash, D. (2007). *The turnaround challenge*. Mass Insight Education and Research Institute.
- Clay, M. M. (1993). *Reading recovery. A guidebook for teachers in training*. Heinemann.
- Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the US population: 2014 to 2060, current population reports* (pp. 25–1143). US Census Bureau.

- Comer, J. P. (Ed.). (1996). *Rallying the whole village: The Comer process for reforming education*. Teachers College Press.
- Cunningham, W. G., & Sanzo, T. D. (2002). Is high-stakes testing harming lower socioeconomic status schools? *NASSP Bulletin*, 86(631), 62–75.
- Darden/Curry Partnership for Leaders in Education. <https://curry.virginia.edu/faculty-research/centers-labs-projects/research-projects/dardencurry-partnership-leaders-education>. Retrieved February 2019.
- Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education*, 61(1–2), 35–47.
- Datausa. (2018). <https://datausa.io/>. Retrieved January 2019.
- Datnow, A. (2006). Comments on Michael Fullan's, "The future of educational change: System thinkers in action". *Journal of Educational Change*, 7(3), 133.
- Decker, S. L. (2008). Intervention psychometrics: Using norm-referenced methods for treatment planning and monitoring. *Assessment for Effective Intervention*, 34(1), 52–61.
- Dewey, J. (1897a). My pedagogic creed. In *Curriculum Studies Reader* (2nd ed.). Routledge.
- Dewey, J. (1897b). The psychological aspect of the school curriculum. In J. A. Boydston & F. Bowers (Eds.), *The early works of John Dewey 1882–1898: Vol. 5. 1895–1898*. Southern Illinois University Press. (Original work published in 1897).
- Dewey, J. (1900). *The school and society*. University of Chicago Press.
- Dewey, J. (1916). *Democracy in education*. The Free Press.
- Dewey, J. (1938). *Experience and education*. Collier Books.
- Dewey, J. (2000). My pedagogic creed (1897). *Philosophical Documents in Education*, 2, 92–100.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(3), 15–18.
- Elmore, R. F. (2004). *School reform from the inside out: Policy, practice, and performance*. Harvard Education Press.
- Fullan, M. (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10(2–3), 101–113.
- Hargreaves, A. (1982). The rhetoric of school-centred innovation. *Journal of Curriculum Studies*, 14(3), 251–266.
- Harriman, D. *University of Virginia School turnaround report*. http://www.darden.virginia.edu/web/uploadedFiles/Darden/Darden_Curry_PLE/UVA_School_Turnaround/UVASTSPAnnualReport2008_Except.pdf. Retrieved February 7, 2019.
- Harris, A., & Chrispeels, J. (Eds.). (2008). *International perspectives on school improvement*. Routledge.
- Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110.
- Hood, J., & Ahmed-Ullah, N. (2012, February 6). School reform organization gets average grades. *Chicago Tribune Newspaper*.
- Hopkins, D. (2011). *Powerful learning: Taking educational reform to scale*. Education Policy and Research Division, Office for Policy, Research and Innovation, Department of Education and Early Childhood Development.
- Hopkins, D. (2013). Instructional leadership and school improvement. In *Effective leadership for school improvement* (pp. 65–81). Routledge.
- Johnson, R. K. (1997). The Hong Kong education system. In *Immersion education: International perspectives* (p. 171). Cambridge University Press.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. The Wallace Foundation.
- LeTendre, G. K., Baker, D. P., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teachers' work: Institutional isomorphism and cultural variation in the US, Germany, and Japan. *Educational Researcher*, 30(6), 3–15.
- Levin, H. M. (1987). Accelerated schools for disadvantaged students. *Educational Leadership*, 44(6), 19–21.

- Levin, H. M. (1998). Accelerated schools: A decade of evolution. In *International handbook of educational change* (pp. 807–830). Springer.
- Lewin, K. (1947). Group decision and social change. *Readings in social psychology*, 3(1), 197–211.
- Mass Insight Education. (2012). *The turnaround challenge*. <http://www.massinsight.org/stg/research/>. Retrieved April 16, 2019.
- McDermott, K. A., & Jensen, L. S. (2005). Dubious sovereignty: Federal conditions of aid and the no child left behind act. *Peabody Journal of Education*, 80(2), 39–56.
- McGregor, D. (1960). Theory X and theory Y. *Organization Theory*, 358, 374.
- McLoughlin, J. (1990). *The demographic revolution*. Jane Hardback Book.
- Meyer, J. W. (1977). The effects of education as an institution. *American Journal of Sociology*, 83(1), 55–77.
- Miles, R. E., Jr. (1967). The case for a Federal Department of Education. *Public Administration Review*, 27(1), 1–9.
- Morrison, K. (2012). *Causation in educational research*. Routledge.
- Nunnery, J. A. (1998). Reform Ideology and the locus of development problem in educational restructuring: Enduring lessons from studies of educational innovation. *Education and Urban Society*, 30(3), 277–295.
- Perry, T., Moses, R. P., Cortes, E., Jr., Delpit, L., & Wynne, J. T. (2010). *Quality education as a constitutional right: Creating a grassroots movement to transform public schools*. Beacon Press.
- Pew Research Center. (2015). *Population projections*. Retrieved at <https://www.pewresearch.org/topics/population-projections/2015/>
- Pilton, J. (2009). *Policy convergence in education*. Unpublished manuscript, Lehigh University.
- Reardon, S. F., Robinson, J. P., & Weathers, E. S. (2008). Patterns and trends in racial/ethnic and socioeconomic academic achievement gaps. In *Handbook of research in education finance and policy* (pp. 497–516). Routledge.
- Reynolds, C. R., Livingston, R. B., Willson, V. L., & Willson, V. (2010). *Measurement and assessment in education*. Pearson Education International.
- Slavin, R. E. (2008). Perspectives in evidence-based research in education – What works? Issues in synthesizing educational program evaluations. *Educational Researcher*, 37, 5–14.
- Slavin, R. E. (2010). Experimental studies in education. In *Methodological advances in educational effectiveness research* (pp. 116–128). Routledge.
- Slavin, R. E., Wasik, B. A., & Madden, N. A. (1996). *Every child, every school: Success for all*. Corwin Press.
- South Carolina Department of Education. (2018). Retrieved at <https://ed.sc.gov/>
- Steiner-Khamisi, G. (2006). The economics of policy borrowing and lending: A study of late adopters. *Oxford Review of Education*, 32(5), 665–678.
- Stepler, R., & López, M. H. (2016). *US Latino population growth and dispersion has slowed since onset of the great recession*. Pew Research Center [Internet].
- Stringfield, S. C., & Yakimowski-Srebnick, M. E. (2005). Promise, progress, problems, and paradoxes of three phases of accountability: A longitudinal case study of the Baltimore City public schools. *American Educational Research Journal*, 42(1), 43–75.
- The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. The National Commission on Excellence in Education.
- U.S. Census Bureau. (2011). Retrieved at <https://www.census.gov/newsroom/releases/archives/population/cb12-tps83.html>.
- U.S. Census Bureau. (2018a). Retrieved at <https://www.census.gov/programs-surveys/popest.html>.
- U.S. Census Bureau. (2018b). Retrieved at <https://www.census.gov/topics/income-poverty/pov-erty.html>.
- U.S. Department of Education. (2016). *Every Student Succeeds Act (ESSA)*. Retrieved from <https://www.ed.gov/essa?src=rm>, January 2020.
- Wiseman, A. W. (2010). The uses of evidence for educational policymaking: Global contexts and international trends. *Review of Research in Education*, 34, 1–24.
- Ylimaki, R. M., Brunderman, L., & Moyi, P. (2019). Balancing evidence-based and humanistic education in school development for underperforming Arizona and South Carolina schools. *Leadership and Policy in Schools*. <https://doi.org/10.1080/15700763.2019.1695850>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.