Teachers or Programs?

A Historical Perspective on Where Trust Is Placed in Teaching Reading

James V. Hoffman and P. David Pearson

Trust lies at the heart of a successful relationship between professionals and those they serve. When strong levels of trust exist, professionals are able to make decisions that are informed and responsive to individual needs. Teaching, as a profession, is no exception. Trust is necessary for teaching to be effective, but trust in teachers as professionals is not common—which contributes to instructional practices and student achievement levels that are below expectations. It is not parental trust of teachers that is in question; in fact, parental trust in teachers is quite high (Mendes, 2010). It is within the educational system where trust in teachers—the trust that allows teachers to practice their profession—falls short. Nowhere are the complex trends surrounding trust within the teaching profession more clearly revealed than in the teaching of reading. Trust in the teaching of reading from a historical perspective has typically been framed as a choice between trusting teachers or trusting programs. Should teachers be trusted to make important and consequential decisions about their students, or should we trust the programs and materials used shape practices and insist that teachers follow guidelines and teachers' manuals faithfully? In this chapter, we will explore moments in the history of teaching of reading that reveal a lack of trust in teachers as professional decisionmakers and an equally deferential faith that programs and materials can solve the challenges faced in the teaching of literacy. We will argue that, within the educational system, trust in teachers to make decisions and shape a curriculum that is responsive to learners and contexts has been soft in comparison to trust in programs. The current movement toward the implementation of the Common Core State Standards may be one more step toward trusting programs instead of teachers. We will argue that this imbalance in trusting programs/materials over teachers is not in the interest of learners. In the end, we will suggest an alternative future that has the potential to transform reading programs and materials from mechanisms of technical control into tools that serve as resources for trusted, professional teachers to enact effective instruction.

We have adopted a historical lens for the treatment of this topic, with a particular focus on the past 50 years in the United States—the very period adopted by the authors of this volume to characterize trends in reading theory and practice (see Preface, this volume). We will analyze this question of trust across five periods of educational practice

leading up to the present. We label these periods Utility, Accessibility, Efficiency, Effectiveness, and Accountability. In adopting this historical framework, we suggest something of a grand narrative for the imbalance in levels of trust we have identified. We document this imbalance by focusing on moments of disruption or discontinuity in our history that are revealing forces that shape institutional policies. The time periods we offer are, at best, rough estimates that are tied not to particular historical events but to our reading of the broader forces that have shaped education more generally and the particular approaches to materials and programs that prevailed in each period. We hope the treatment of these periods will serve to ground the implicit appraisals of teacher and program trust.

UTILITY: 1630–1830

In colonial America, schools were formed out of social needs and cultural practices. Monaghan (2005) has carefully documented the emergence of schools as institutions and the qualities of reading instruction in America during the colonial period. Reading instruction was largely shaped by local circumstances, especially the dominant influences of local religious authorities. The nation was mostly rural, and access to schools was very limited. Individual families were left to their own resources in planning for the education of their children. "Dame schools" or home schooling were often the only options for instruction. Because the goals for reading instruction were limited—firmly focused on the spiritual, with some attention to the economic and social (e.g., contracts, laws)—the demand and resources needed for reading programs in schools were also limited (Smith, 1965). Teachers were seldom prepared for their role beyond their own experiences and expertise in reading and sometimes writing. Given its very limited goals (e.g., an ability to read the Bible), reading instruction tended to focus on the technical dimensions of reading. According to Smith (1965), instruction relied on restricted materials (e.g., hornbooks and primers) and limited methodologies (e.g., repetition, memorization, and recitation).

In 1647, a law was passed in the colony of Massachusetts that related to the responsibilities of communities to educate youth and in particular to teach children to read. This law became known as the Old Deluder Satan Law. The law described ignorance as a "satanic ill" to be addressed through education so that "ye ould deluder, Satan, could not use illiteracy to keepe men from the knowledge of ye Scriptures." The law required every town with more than 50 families to hire a teacher and towns with more than 100 families to establish a grammar school (for boys only). These kinds of laws were also passed in other colonies to ensure that children were being taught to read following "The Ordinary Road" from the hornbook, to the primer, to the Psalter, to the New Testament, and finally to the entire Bible (Monaghan, 2005). Monaghan points to this period and these laws specifically as the first steps toward the institutionalization of public education as a responsibility of the government. With this shift came the responsibility for government to both shape the goals and monitor the quality of education.

As goals began to expand in the revolutionary period (toward citizenship and moral character), shifts in instruction and programs soon followed (Smith, 1965). Noah Webster (1797) was a key figure in this period, developing a three-part instruction series: *A Grammatical Institute of the English Language.* Webster's "Blue-Back Speller" was the most popular of the books in this series. Webster's attention to spelling and to the use of a phonic method (tied to the teaching of letter-sound relationships) would challenge the

dominance of the "alphabet spelling" method (tied to letter naming and spelling) and would radically reshape reading instruction because there was a more transparent relation between the letter and its sound in helping students unlock new words.

Perhaps even more important was Webster's attention to stories that exemplified moral behavior, for this would mark the beginning of a movement away from the dominant religious function of instructional text. All of these shifts that impacted practice were tied to institutional and societal influences. Webster (1848) was less concerned about pedagogy than about the standardization of oral and written English that was uniquely American. Letter-sound work would support Webster's goal of standard pronunciation in a country that was divided by dialects. Standardized spelling would ensure a common written language; thus, Webster's lifework on the dictionary of the English language. The focus on moral stories reflected his strong belief in the future of democracy and the forces that would bind the nation rather than divide it. Moral character became a safe middle ground that would appeal to the diverse array of religions in the country, while avoiding the question of whose religious beliefs would take precedence in whatever materials might be developed (Bynack, 1984; Kendall, 2011; Rollins, 1980). In a sense, this new tradition of materials for beginning reading, like so much in the rest of the new nation, championed the separation of church and state.

From its infancy, American reading instruction has adapted to shifts in utility. Shifts in practice followed shifts in societal goals, needs, and structures. Schools and practices were shaped by these shifts and relied on commercial reading materials and a "teacher" who had no formal preparation to teach. Crises, as in the moral imperative to learn to read the Bible for salvation or the need to standardize English pronunciation for the preservation of the nation, were used as leverage to reshape curriculum and teaching. It was widely believed that anyone who possessed strong moral character and the ability to read could be trusted to transfer these skills to others using the materials that embodied core values around language, religion, and citizenship. In this manner, schools (or, perhaps more aptly, schooling) would play a critical role in promoting religious homogeneity, economic prosperity, and national unity. At least until the awareness of a need to expand access to literacy became a part of the American conscience, utility held sway in shaping the teaching of reading. And teachers were the purveyors, not the shapers, of that utility; they did society's bidding.

ACCESS: 1830–1890

The period of education between the late 1830s through the 1880s is often characterized as the period of the Common School Movement (Kaestle, 1983). The population of the United States was growing rapidly largely, as a result of waves of immigration, and there was an accompanying expansion of territory to the west. With this rapid growth came a perceived need to extend literacy both for an expanding range of occupations and because of a desire to extend the rights of citizenship to a broader segment of (the White male) society; thus, the demands for schooling increased dramatically.

The Common School Movement was famously led by politician and educational reformer Horace Mann and focused on the goal of providing free, compulsory, and public primary education for all citizens (Cremin, 1957). Mann, from his leadership position in the state of Massachusetts, aggressively modeled the ways in which state governments

could promote education for all citizens. He believed education was essential for instilling a spirit of democracy and hope in all citizens for a better life. Mann was concerned that educating only the elite would not bode well for the future of democracy. He founded *The Common School Journal* in 1838 and used his widely circulated annual reports on the state of education as a platform to raise attention to the inadequacies of schools in the state of Massachusetts around a range of issues from disdain for corporal punishment and punitive teaching to advocacy for graded schools that would meet the needs of greater numbers of students and school libraries (Filler, 1965).

Influenced by the reforms in education that were taking place in Europe, Mann advocated for reading instruction that focused on the word method, to replace the dreaded drill of alphabetic methods. A flurry of reading materials based on the word method, beginning with Ward's rational method in 1896, competed in the marketplace. McGuffey's readers, relying on the word method, appeared and gained prominence throughout the country (Ruggles, 1950).

As with the previous period, shifts in practice were influenced by institutional changes (within state educational systems, in particular), creating a new set of tools to shape educational practices through public policy at both the state and district levels. From the Common School Movement came the first significant attention to the preparation of teachers, but the teacher's role in shaping the curriculum was not clear. Mann valued investment in teacher preparation, but his work suggests that he envisioned a role for teachers as deliverers of a curriculum rather than as initiators or innovators. The training—not the education—of teachers was emphasized (see Hoffman & Pearson, 2000, for more on this distinction). There was an envisioned role for teachers revealed, in particular, in the positioning of women as elementary teachers by the leaders of the Common School Movement.

Preston (1989) offers a critical analysis of the work of teachers as constructed by leaders in the Common School Movement, with a particular focus on the feminization of teaching. Mann, for example, during his tenure as the Massachusetts secretary of education from 1837 to 1841, put forward numerous arguments for the employment of female teachers, including the curious point that women were superior to men as teachers because of the qualities they possessed "by nature": Mild and gentle manners were considered consistent with "true women's" docility. The feminization of the elementary teaching workforce that took place during this time period is complex and is not the focus of this chapter. However, it is important to note this major disruption in teaching and in society as women moved into professional roles that were then constructed as being docile and nonintellectual.

Again, as in the previous period, we see the use of crisis rhetoric to prompt public action. In the Common School Movement, the crisis was the lack of education for the growing masses, guided by the assumption that ignorance could potentially destroy a democratic society. Whom did our society trust? Teachers, but with qualifications: Society could trust those who were trained in what and how to teach—those who would do what was expected and do it well.

EFFICIENCY: 1890–1970

The Progressive Era in American history was a time of rapid change, social activism, and economic expansion (Buenker, Burnham, & Crunden, 1977; Flanagan, 2007; Gould, 1974). Historians typically frame this era as somewhere between the 1890s through the

end of the 1920s and the start of the Great Depression. The "Efficiency Movement," or the "Economy of Time" reform effort, was a significant part of the Progressive Era (Tyack & Cuban, 1995). The Efficiency Movement emphasized the application of scientific principles to the accomplishment of human work. Efficiency became a centering theme for the Progressive Era. The assembly line, used in mass production, was a driving metaphor for the movement. Determine the desired characteristics of a product and then find the way to produce the product with the smallest allocation of resource costs—both human and material. The "stopwatch" (Taylor, 1913) became the tool that symbolized the movement. Scrutiny over the efficiency of schools in response to rapid population growth soon appeared.

Joseph Rice (1893), a pediatrician, provides a rich example of this fascination with efficiency. He conducted a year-long study of education in school systems across the country during this period and published his findings in serial form in magazine articles and later as a book. Rice employed both observation and student assessments as tools in his inquiry into schools and teaching practices. Although the results of his inquiry are complex, a few generalizations can be offered. First, Rice claimed that, with few exceptions, schools across the country were failing in their responsibility to educate American youth. Second, Rice found that the quality of the teaching in most schools was poor and disorganized and that teachers were largely unsupervised and unsupported. Rice saved his harshest criticism for leadership (at the school and district levels). The failure of the system should not be blamed on the teachers, Rice argued, but on the system that has failed them and their students. Rice (1913) called for massive reforms in school organization and leadership that echoed the principles of the Efficiency Movement for clear goals, measurable outcomes, and close supervision and support.

The Efficiency Movement, however, was not just about managing resources. It was accompanied by a particular vision of society. Just as everyone has a place in the manufacturing of a product (e.g., workers, managers, owners), everyone has a place and role in society. Do your part. Accept your role. Both individuals and the society would thrive as a result. The capacity of individuals to fill certain roles was not distributed equally. Some individuals are suited for menial tasks or physical labor, others for intellectual work, and still others for creative endeavors. Schools should prepare individuals for their role in life—and the sooner a given individual's future role could be ascertained, the sooner schools could get him or her into an educational track that would offer apt preparation for pursuing that role. Less education, and education focused on the technical and vocational, would suit the less intelligent, those who were destined for factory roles and service occupations. Higher levels of education would be required for the more intelligent, who were bound for professional roles and responsibilities. This sorting would take place at an early stage in educational systems based on scientific assessments of intelligence and aptitude. Over the full range of the 20th century, tracking, grouping, and limiting access to schools with high standards would serve as the mechanisms to achieve this end (Oakes, 2005).

William S. Gray, commonly regarded as the father of reading education, was active in this period. Like many of his colleagues, he advocated for the scientific management of schools through testing. He guided periodic reviews of schoolwide reading programs to assess their quality. Standardized measures, with measures of reading rate prominent, were used to assess outcomes, and the system was examined for the efficient and scientific management of resources toward outcomes. The "whole-word" or

"look-say" method, distinguishable from the word method in the previous period that was focused on meaning, emphasized rapid and automatic word recognition through repeated exposure. Speed, efficiency, and practice mattered (Carver, 1990). In the quest for efficiency there was a strong push toward standardization of reading programs during this time through the mechanism of state laws and regulations for the purchasing of core textbooks across the country. During the 1920s, there was a dramatic rise in attention to state textbook adoption policies that promised free textbooks to schools (Tidwell, 1928). Though there were many motivations for state textbook adoptions (e.g., cost savings or guarantees of equal access to good pedagogy across districts within the state), the primary goal was standardization and control over the curriculum (Tulley & Farr, 1985). Basal readers fit neatly into this scheme of state procurement and control. From the publishers' viewpoint, this movement toward state adoptions offered the promise of reward and a fairly stable target for content. Basal reader systems became identified with efficient and organized programs that would promote learning to read. Gray et al.'s own Scott Foresman basal series, the Curriculum Foundation series (1940–1948; 1951–1957) featuring Sally, Dick, and Jane, became the standard for the field as an organized reading program.

In line with the goal of efficiency and sorting, the reading community began to promote a view of reading programs as serving different kinds of students in different ways. A school reading program consisted of a developmental program (designed for the majority of the children—"normal"), the corrective program (for children who were falling behind—with a shared responsibility for the classroom teacher and the reading specialist), and the remedial program (for children who were very behind and required the support of a reading specialist in one-on-one settings). Occasionally—but only occasionally— there might be commercial programs that paralleled these classifications, and sometimes even programs for gifted students as well. For example, in the 1960s, Scott Foresman published, in addition to a foundational program (the New Basic Readers: Curriculum Foundation Series; Robinson, Monroe, Artley, & Greet, 1960–1962), a remedial program with the label of *Open Highways* (Robinson et al., 1967) and a gifted program entitled *Wide Horizons* (Robinson et al., 1965). Publishers Lyons and Carnahan offered a different twist on this same theme of differentiation. In their 1962 Companion Readers (Bond & Cuddy, 1962), the same story was written in three different versions—one *at* grade level for most students, one *below* grade level for struggling readers, and one *above* grade level for gifted readers. The bet was that a teacher could hold a whole-class discussion around the core ideas in a story because all the students had versions that they actually stood a chance of reading on their own. But the logic of differentiation and getting students placed properly was the same as for the Scott Foresman approach.

The drive toward efficiency can certainly be framed from an economic argument and the scarcity of resources principle. However, it can also be framed from a sociological perspective around schools working toward an envisioned society with emphasis on the American dream that hard work leads to progress and economic well-being when everyone plays their role. As Rice (1913) noted, the crisis (failure) in schools was not the fault of teachers but the fault of administrators who failed to apply proper organizational structures and guidance. Where was trust placed during this era? In scientific management that would guide teachers to do their job—with materials defining the track they should follow.

EFFECTIVENESS: 1965–1985

In her classic book, *Learning to Read: The Great Debate*, Jeanne Chall (1967) described the period of reading instruction from the turn of the 20th century up through the 1960s as a period of gathering consensus around practice. This consensus was embodied in the leading basal reading programs during this time, with a focus on readiness, leveled texts, word selection based on a meaning-frequency principle, sight-word ("look-say") methods with supporting word repetition in the readers, and ability grouping within the classroom. There was widespread agreement throughout the first half of the 20th century that reading goals were being met through this consensus approach. This view was disrupted by the scrutiny placed on schools as the "Sputnik era" emerged and questions regarding effectiveness and equity bolted onto the scene. Concurrent with the Sputnik wake-up call were national reports, such as the Coleman Report (J. S. Coleman et. al., 1966), documenting patterns of achievement in reading and other areas that showed large and growing discrepancies between White and African American youth, and Flesch's (1955) scathing critique in *Why Johnny Can't Read*, decrying the very same methods that Chall documented as prevalent in her 1967 book. President Johnson, through his Great Society programs, poured money into federal programs and research to address issues of effectiveness. This period marked the first significant involvement of the federal government in education, including shaping educational research, based on civil rights.

During this time period, there were four particular research efforts focused on program issues and the teaching of reading that are important to consider: the Effective Schools studies, the First-Grade studies, the Follow-Through studies, and the Rand Change Agent studies.

Effective Schools Studies

A number of researchers during this period challenged the general characterization of schools as having a minimal influence on achievement based on the findings from the Coleman Report (J. S. Coleman et al., 1966). They questioned the assertion that achievement was determined by qualities (e.g., intelligence and SES) outside the control of schools. These researchers often examined the interaction of achievement and context in outlier schools where school and program characteristics did make a difference (Hoffman & Rutherford, 1984; Purkey & Smith, 1983). Drawing on the work of Weber (1971) in schools in Harlem and other researchers who examined high-success schools in the context of poverty, Edmonds and Frederiksen (1978) documented the qualities of programs (in urban schools in particular) that made a difference despite individual and community characteristics—schools that were "beating the odds" predicted by their demographic makeup. These qualities ranged from an orderly and safe climate conducive to teaching and learning and principal leadership to a clear instructional focus (on reading), high expectations for all learners, and systematic plans for assessment. One strong assertion that came out of this literature is that these effective schools were not places that had gathered together a large number of outstanding teachers. In fact, many of the research reports went out of their way to describe the teachers and teaching as not particularly creative or impressive—in fact, these were described as rather "businesslike" (Rosenshine & Furst, 1971).

First-Grade Studies

A second research effort, specifically focused on reading instruction, took the form of a national comparison study of effective programs supported by the federal government. It was officially known as the Cooperative Research Program in First-Grade Reading Instruction (Bond & Dykstra, 1967) and commonly referred to as the First-Grade studies. Twenty-seven different research projects around the country were coordinated to address three questions:

1. To what extent are various pupil, teacher, class, school, and community characteristics related to pupil achievement in 1st-grade reading and spelling?
2. Which of the many approaches to initial reading instruction produces superior reading and spelling achievement at the end of the 1st grade?
3. Is any program uniquely effective or ineffective for pupils with high or low readiness for reading?

The studies compared various approaches, methods, and materials for the teaching of beginning reading, including traditional basals, language experience, a "linguistic" approach , systematic phonics-based methods, and the "*ita*" (initial teaching alphabet). The questions that guided the study reveal the widely held view of the time that programs (i.e., approaches, materials, and methods) were at the heart of effective teaching. In some sense, the notion that methods and program comparison studies were the focus for the study of effectiveness reflects trust in programs rather than teachers as the critical variable. The fact that none of the programs or approaches showed any great advantage in student outcomes (both student success and failure were documented across all programs) led the researchers to speculate that teacher quality was the variable that was most responsible for achievement differences. Because teachers were not included as a variable in the study and few of the studies even monitored the level of implementation, this assertion was left as nothing more than conjecture.

The Follow-Through Studies

A third research effort that focused on the teaching of basic skills in the early grades was the Follow-Through studies (Egbert, 1981). The Follow-Through Program (1968–1977) was initiated in response to findings from evaluation studies that showed the academic gains realized through the introduction of Head Start programs in schools serving low-income communities seemed to wash out as students entered formal schooling (Maccoby & Zellner, 1970; Stebbins, St. Pierre, Proper, Anderson, & Cerva, 1977). Follow-Through was designed to provide academic support to bridge the movement of students from Head Start into early primary education. Rather than specifying a particular program for implementation and using a randomized-control model for evaluation, the federal government gave local school districts the choice to select from a range of programmatic options. The purpose was to increase community control and investment in the outcomes (Egbert, 1981). The included programs were as diverse in their philosophy and execution as Direct Instruction (DISTAR) and the Behavioral Analysis Model, which anchored the scripted-structured end of the programmatic continuum, to the Bank Street Model and Open Education Model, which were the least constrained and structured of the

programs. Most of the program models, outside of DISTAR and the Behavioral Analysis Model, were rather vague, underdeveloped, and barely field-tested (Elmore, 1977).

The intervention and research model for the Follow-Through studies is referred to as planned variation. Stallings took a leadership role in evaluating the effects of the Follow-Through Program and the features of these programs that were associated with positive learning outcomes (Stallings, 1976; Stallings & Kaskowitz, 1974). As with the First-Grade studies, there was no clear winner; rather, there was variation across sites and programs in implementation and success. In general, the research revealed that the features associated with effective programs included:

- a clear focus on academic outcomes and academic engaged time;
- time spent working in textbooks and academic workbooks (as opposed to time spent with puzzles, games, toys, and the like);
- grouping and whole-class instruction;
- high program structure and sequence;
- carefully prescribed teaching practices; and
- academic-focused feedback (acknowledgment, praise, and positive and negative corrective feedback).

Overall, there was general support for direct instruction with clearly specified learning outcomes. Although the results and claims surrounding Follow-Through were widely questioned (see Elmore, 1977; House, Glass, McLean, & Walker, 1978), the findings continued to be used for at least another 3 decades to support program effectiveness in reading that positions teachers in a delivery role (Carnine, Silbert, Kame'enui, & Tarver, 2009).

The Rand Change Agent Studies

A fourth research effort focused on reform at the programmatic level related to the federal government's support for the creation and transfer of effective programs. The Elementary and Secondary Education Act (in particular, the Education Amendments of 1974) provided support for innovative and exemplary projects that provide creative or imaginative solutions to problems in curriculum and teaching. Some of these grants were designed to support development work, while other grants were designed to support institutions working to adopt programs that had been created through development grants. The National Diffusion Network (NDN; 1974–1995) was the first federally sponsored effort to identify and spread innovative educational programs. The NDN would conduct an evaluation of the data on the effectiveness of innovative programs and certify those that passed the demonstration of effectiveness for diffusion. The program was administered through the Office of Education and was designed to make use of the best ideas from the innovative programs. A large number of these programs focused on improving reading achievement.

The federal government sponsored an independent evaluation of this process, designed to spread innovation. The Rand Corporation was contracted to study the effective transport of innovative programs from one site to the next with a particular focus on replication of the outcomes related to student achievement. Paul Bermann and Milbrey McLaughlin directed this study, which was eventually reported as the Rand Change Agent studies, undertaken from 1973 to 1978 (McLaughlin, 1976, 1990). It was built on

a framework for change constructed by Berman and McLaughlin (1976). They examined four federal programs and 293 projects in 18 states. Overall, they found the level of implementation of programs to be very low and uneven. They identified a set of characteristics that seemed to be very ineffective in supporting program adoption and a set of strategies that they associated with success.

Some factors associated with limited implementation:

- Reliance on outside consultants
- Packaged management approaches
- One-shot, pre-implementation training
- Pay for training
- Formal, summative evaluation
- Comprehensive systemwide projects

Some strategies associated with effective implementation:

- Concrete, teacher-specific, and extended training
- Classroom assistance from local staff
- Teacher observation of similar projects in other locations
- Regular project meetings that focused on practice
- Teacher participation in project decisions
- Local development of project materials
- Principals' participation in training

This report concluded that effective change in schools does take place through adoption, but also through mutual adaptation: the adaptation of a project or policy and the organizational setting to each other. Too much emphasis on fidelity to the original program could undermine the entire effort to effect positive programmatic changes that would lead to increased student achievement (Elmore & McLaughlin, 1988).

The effectiveness era was characterized by a belief that research could provide answers to the important questions surrounding teaching and literacy achievement—in particular, the disparity of achievement related to socioeconomic factors—through the careful study of programs (i.e., approaches, materials, and sometimes training) as they are implemented and evaluated. Despite the mixed results and the failure to demonstrate transfer, faith remained strong that the answers to questions of equity would be found in trust at the program level and the training of teachers who could follow programs with a high degree of fidelity. Who was trusted during this period? Not teachers. "Teacher-proofing" the curriculum became the ideal.

ACCOUNTABILITY: 1985–????

The start of the accountability movement in education is often associated with the publishing of the report *A Nation at Risk* (National Commission on Excellence in Education,

1983). Echoing the Sputnik era critics of the previous period, the authors and sponsors of the *Nation at Risk* report claimed that educational system in the United States was failing its citizens to such an extent that it had become a matter of national security. The document, spearheaded by Secretary of Education Terrel Bell, advised educational systems to model themselves after businesses, including the advice to hold people accountable for results in return for the resources being allocated. In particular, the report called for the application of more rigorous and measurable standards—and the report explicitly cited the success of the measurable standards movement in business and industry.

The tools for reform and accountability evolved in the years following *A Nation at Risk*. The first big enhancement came at the end of the 1980s, with the influential Charlottesville National Governors Conference in 1989 (Bill Clinton, then governor of Arkansas, convened the conference). Out of that meeting came all of the apparatus to encourage all disciplines to follow the lead of the National Council of Teachers of Mathematics (NCTM) in developing highly rigorous and specific *content* standards (specifying what students should know and be able to do), as well as standards for *assessment* (what cut scores on what tests would tell us whether individual students or teachers or schools were achieving the content standards?). (See National Research Council, 1999, for a compelling account of the evolution of the standards movement.) The tools of standards-based reform were standards, assessments to measure their achievement, and the stakes or consequences that come with an accountability system. These consequences took the form of rewards or sanctions for good or poor performance, particularly at the school level, but there were—and are—examples where the onus falls at the student level—for example, with state and district retention policies.

Some progress in implementing these new policy tools was achieved during Clinton's presidential years, mainly through the auspices of the Improving America's Schools Act of 1994; however, these various tools were not finally brought together legislatively until the No Child Left Behind Act of 2001. This act established an accountability system for states, school districts, and schools receiving federal funding. The law required states to establish academic standards, tests to measure their mastery, and an accountability architecture that would require schools to make annual progress toward having every student achieve the standards until, in 2014, all students in every school would achieve the mastery cut score—a policy version of "Lake Wobegone," where all children are above average! Schools also had to make annual progress toward closing the achievement gaps between various demographic groups, particularly groupings based on race, language, income, or intellectual status. The law also required states to identify schools and school districts that were not making adequate yearly progress and to mandate that these low-performing schools follow a step-by-step process for either turning themselves around or risk closure. Although neither NCLB nor its reading section, Reading First, was a "program," the context of these mandates shaped programs in clear and powerful ways. The accountability legacy of this era for reading can be characterized by our version of the "big five" in literacy policy: standards, stakes, scripts, sorts, and supervision.

Standards

The standards movement that began in the late 1980s has been gaining momentum ever since. States such as Texas have spirited the movement with variations that began

as "opportunity standards" around curriculum areas and have moved toward performance standards that are mostly grade-level bound. We have already alluded to the two lynchpins of the standards movement—content standards (what students know and can do) and performance standards (what scores on which tests indicate whether they have achieved standards). But there was the third key standard that, while part of the original conceptualization of standards, has lost its punch since the early days of the standards movement—namely, opportunity to learn standards. They were supposed to be the "quid pro quo," the resources that schools, districts, and states would offer in return for schools and teachers holding themselves and their students to account for the content and performance standards. Alas, they seem to have dropped out of the standards equation, so what we are left with is teachers, schools, and students being held accountable for particular standards without a reciprocal allocation of curricular and material resources to achieve those standards.

Stakes

The rise of high-stakes assessments across the country was (and continues to be) fast and furious (see critiques by Linn, 2000, and Nichols, 2012). These high-stakes assessments typically take a census approach (everyone gets assessed on the very same test) rather than a sampling approach (samples of students get assessed on different facets of the learning goals and overall performance at the school or district level on each of these facets is inferred from the performance of the samples). The choice of a census model is certainly not driven by economics (a state report based on sampling could be conducted on a much wider range of reading standards at a fraction of the cost of the current census model), but rather by the intent to shape compliance with the standards through the logic of accountability (i.e., make the stakes for low performance high for everyone, including students!). It should be noted that this "every student takes the very same test" approach to state-level testing comes at a very high cost, not only in terms of the actual dollars spent but also in terms of content covered by the test: Only a very small sample of all the important curricular outcomes can be assessed in the hour or 2 allocated for a reading or a writing assessment. In the case of reading comprehension, this means that only a very small sample of passages can be used from which to draw a conclusion about students' collective performance.

This accountability mechanism allows stakes to affect players at all levels in the system. Students' lives are affected in terms of promotion, graduation, immediate rewards, or placement in mandatory tutoring programs. Teachers are subject to bonus pay, contract probation, or even termination. Schools get labeled on a continuum from exemplary to failing, threatened with closure, and subjected to involuntary faculty reassignment. Even communities and neighborhoods suffer or benefit from the impact of school test scores on property values. The high-stakes assessments are tied to the standards in a structure that is criterion-referenced, although the scores are often interpreted on normative grounds (e.g., what percentage of students should achieve the magical mastery cut score).

Scripts

The third piece of the accountability puzzle takes the form of scripts, where the script metaphor is intended to capture the concept of program fidelity. The No Child Left Behind Act required program resources to be used only to support teaching strategies and

materials that had been shown effective through scientific research. These commercial programs suggested that the teacher's role was to follow the manual instructions without deviation. Scripted programs were not new phenomena in the 2000s. In fact, they are as old as basal programs themselves, which have, at least since the 1920s, offered teachers explicit directions for how to teach skills. For example, the DISTAR (Bereiter & Engleman, 1966) program of the 1960s championed the notion of scripted instruction as reflecting research on effective teaching, along with a trust in materials and practices. But the surge in popularity and use of scripted programs through NCLB and Reading First was unprecedented. Although scripts were found only in a few programs during the Effectiveness period, now scripts are found in the most popular core reading materials (such as the McGraw-Hill/SRA version of Direct Instruction). In fact, state adoption policies called for districts to ensure that teachers followed the scripts and pacing guides that align with the standards and stakes.

Sorts

In the Accountability era, attention to grouping and tracking operates in ways that still reflect the efforts initiated in the Efficiency period. But some of the grouping plans and tracking have resulted from ad hoc efforts in response to the pressures of high-stakes assessments. "Bubble Kids" or "Triage for Testing" have become a common part of the discourse in programs across the country. Because most high-stakes tests are criterion based, passing (i.e., reaching the magical cut score) or not passing is the only thing that matters. Students who are going to pass the test regardless of the instruction they receive become a low priority for resources. Likewise, students who will *not* pass the test regardless of the support offered also become a low priority for instructional resources. "Bubble Kids," or those who may be able to pass with intense support, receive the bulk of the attention—all directed at the targeted outcome (i.e., achieving a passing score on the test).

As with instructional programs, during the reign of NCLB/Reading First, assessment tools underwent an official review by a federally funded technical assistance center. And assessment tools that survived the review process, including classroom assessments designed to monitor progress and diagnose individual needs, were strongly recommended (with the moral force of mandate) for use by districts and schools receiving NCLB funds. No assessment tool received more uptake for Reading First than DIBELS (Good & Jefferson, 1998; Good & Kaminski, 2002), at least in part because of its blessing as an officially validated tool. Even through DIBELS was designed as a progress monitoring tool (taking a reading on student performance at regular intervals), it became widely used as a diagnostic tool used to shape instruction in school settings. The net effect of DIBELS use was to focus corrective and remedial efforts on the low-level fundamental skills of phonemic awareness and phonics that it measured, thus relegating low-performing students to a steady diet of basic skills in their reading programs, with no opportunity to engage in advanced comprehension or critical thinking.

Supervision

The final policy tool in the NCLB toolkit was supervision, especially supervision enacted as classroom observation to ensure that the scripts and pacing guides were being followed with a high degree of fidelity. Program fidelity has become a key theme in the

accountability period, with the goal of every teacher enacting programs in exactly the same manner. Here again, NCLB and Reading First took the lead in enacting this fifth element by supporting coaches in observing and supporting teachers toward successful implementation. This is not a coaching model that takes a reflective stance toward teacher learning, but rather one that emphasizes compliance and fidelity of implementation.

The Aggregate Effect of These Levers

Taken together, these five elements of the accountability movement, particularly as the movement reached its acme in the NCLB era, have come to define what counts as a reading program. Alignment is seen as the key process in program development. Align your goals, the materials and assessments you use, the professional development you receive, and the guidance you get from coaches in your classroom. Sort students into tracks using assessment tools. Provide scripts and pacing guides for teachers to follow, and supervise implementation according to the goals of the program. Trust the program and insist that teachers follow the plan—that's the formula for success. Who is trusted in such a system? Programs—not teachers. The freedom for teachers to become responsive to learners had reached a low point in history as Obama took office and the era of the CCSS began. But has anything really changed since the CCSS came into being? That is the topic to which we now turn.

THE ROLE OF READING PROGRAMS IN THE COMMON CORE STATE STANDARDS

It is hard to characterize the post-NCLB era ushered in by the one-two punch of the Obama administration's Race to the Top initiative (U.S. Department of Education, 2009) and the appearance of Common Core State Standards (National Governors Association Center for Best Practices [NGA Center] & Council of the Chief State School Officers [CCSSO], 2010). One reason why it is difficult to characterize the movement is that these initiatives are at once both a critique of the NCLB Bush policies and a continuation of them. What continued from the Bush years was the accountability apparatus—standards for what students should know and be able to do; assessments to measure progress toward achieving them at the school, district, and state level; and rewards or sanctions for achieving or failing to achieve the standards. The critique focused on the standards and the assessments to measure them, implying, of course, that it was not the accountability construct that was the problem in NCLB but rather the fact that our curriculum and pedagogy were being guided by flawed, low-level standards and misguided assessments. The remedy, according to the logic of the Obama administration, was better (i.e., higher and more rigorous) standards and better (i.e., more challenging) assessments to measure students' mastery of the higher standards. The promise of the CCSS was to bring rigorous curriculum to all students.

This current logic hearkens back to the beginning of the standards movement in the early 1990s, when the motto was something like "If we are going to teach to the test, then let's have tests worth teaching to" (Resnick, 1993). In the 1990s, this logic drove the development of performance assessments (e.g., the California Learning Assessment System, 1994). Such tasks required students to respond to complex, challenging, multistage tasks that emphasized integrating knowledge and insights across multiple sources and using

that knowledge to address some sort of pressing complex social, scientific, or literary issue, such as whether global warming is real, whether Martin Luther King's reliance on the nonviolent strategies of Gandhi served the interests of civil rights reform in the United States, or how O. Henry achieved irony in his short stories. In 2010, this same logic drove the Obama administration to invest over $365 million in two national consortia, Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (SBAC), to develop assessments that would do justice to the higher standards mandated by the CCSS.

The CCSS, which have been voluntarily adopted by each of the 44 participating states (making them a state and not federal initiative), promise a rigorous curriculum for students of precisely the sort that will earn them entry into higher education and/or a solid, secure career in the workforce. In the bargain, they will be prepared to do well on the new tests that are being developed, somewhat ironically, with federal support. The vision of student excellence is exemplary, almost inspiring:

> Students who meet the Standards readily undertake the close, attentive reading that is at the heart of understanding and enjoying complex works of literature. They habitually perform the critical reading necessary to pick carefully through the staggering amount of information available today in print and digitally. They actively seek the wide, deep, and thoughtful engagement with high-quality literary and informational texts that builds knowledge, enlarges experience, and broadens worldviews. (NGA Center & CCSSO, 2010, p. 3)

The CCSS developers' view of the role that standards should play in the classroom reflects progressive moral and ethical values about teachers and teaching. The body politic has the right to set the ends or goals for our schools and students, but teachers must have the prerogative to determine the means of achieving those ends:

> By emphasizing required achievements, the Standards leave room for teachers, curriculum developers, and states to determine how those goals should be reached and what additional topics should be addressed. (NGA Center & CCSSO, 2010, p. 4)

Again, in discussing the implementation of criteria for judging the complexity of text, the developers of the CCSS ceded the right to moderate quantitative assessments (e.g., readability scores) of complexity to teacher judgment:

> While the prior two elements of the model focus on the inherent complexity of text, variables specific to particular readers (such as motivation, knowledge, and experiences) and to particular tasks (such as purpose and the complexity of the task assigned and the questions posed) must also be considered when determining whether a text is appropriate for a given student. Such assessments are best made by teachers employing their professional judgment, experience, and knowledge of their students and the subject. (NGA Center & CCSSO, 2010, Appendix B, p. 4)

There is hope in some camps (for glimpses of this guarded optimism, see Pearson, 2013; Pearson & Hiebert, 2013) that the new CCSS will be a game changer in that they will provide a different vision of what it means to meet high and rigorous standards and what it means for states and districts to come through with the right curricular and material

resources to help schools help teachers help students meet the standards. Whether all this will happen remains to be seen. On paper, it appears that in contrast to earlier policies that privileged programs over teacher prerogative, the CCSS are willing to trade teacher accountability to high and rigorous standards for at least a modicum of teacher choice in curricular and pedagogical matters.

We can't yet know whether the standards will be implemented in a manner that will allow them to remain true to these lofty aspirations. There is already reason to believe that some of the documents guiding implementation of the CCSS betray these aspirations. For example, Pearson (2013) has pointed out that the highly influential Publishers' Criteria, written by the very authors who wrote the CCSS-ELA (D. Coleman & Pimentel, 2012), put forward a model of reading comprehension instruction, dubbed "close reading," that undermines the emphasis on interpretation, integration, and critique that characterizes two-thirds (Standards 4–9) of the CCSS for reading. In the same paper, Pearson (2013) points out that the same Publishers' Criteria, because they offer such highly specific instructional guidance, can quickly erode teacher prerogative. That is, when teacher manuals offer fine-grained, step-by-step guidance (very much like the scripts used during the NCLB accountability era), they leave teachers with few choices to make, rendering the promise of prerogative a hollow sham. These possibilities provide room for skepticism on the question of teacher trust as educators begin the arduous process of implementation.

BUILDING TRUST ALL AROUND: A CHALLENGE FOR PROGRAM DEVELOPERS AND A NEW GENERATION OF TEACHERS

Thirty years may be a very short span of time in an historical sense, but 30 years of life with accountability is surely enough in the face of the failure of this movement to serve students well. We look forward to the time when we can put an end date on the Accountability period. We close our chapter with a focus on two shreds of hope that we have identified in the current and somewhat discouraging state of affairs. One lies in the consideration of a movement situated in the very teacher guides that we have decried for their complicity in foisting scripts on teachers to achieve a high level of fidelity in program implementation. Beginning with the influential essay of Ball and Cohen in 1996, several scholars have undertaken the study of "educative materials" for teachers (Davis et al., 2014; Krajcik & Davis, 2005). With all these manuals available to teachers, the fundamental idea is, why can't educators exploit their ubiquitous use to promote deeper and broader teacher knowledge of curriculum and pedagogy, with the proviso that those who design educational materials for teachers undergo an elaborate iterative design process in which evidence is used to refine materials at every step in the development process? The hope—indeed, the expectation—is that when teachers use materials developed in such a process, they actually learn more about their craft. To our knowledge, we have yet to witness an empirical test of the claim that teachers can refine their craft with well-designed materials, but the concept is so eminently sensible that it seems almost compelling.

A second, and perhaps more absorbing, shred of hope comes in considering teachers—in particular, in the consideration of the young teachers who are entering or have just entered the profession. This new generation of teachers is the first to have personally experienced the weight of accountability throughout their lives as students in schools. They know the system from the inside and may, therefore, be the ones who

are best equipped to topple the Accountability era and replace it with a new vision for the teaching profession. We challenge these young professionals to consider the virtue of responsibility as a stance that can lead to meaningful reform. Teachers who take responsibility for their learning, for the quality of their service to children, and for the shaping of their profession would be a formidable force for change and would usher in a new era of Responsibility.

The simple fact is that accountability sets a much lower bar for excellence than responsibility. Educators have decades of evidence that trying to raise the quality of schooling by raising accountability only distorts teachers' efforts to be responsible for their work. There is no evidence that raising accountability standards has, as promised, driven less effective teachers from the classroom. The performance standards associated with accountability reflect an easily rendered mimicking of what real teaching involves. Teachers who take on responsibility for their professional lives would mark a shift toward trust that enables the design of materials and programs that can serve as tools for achieving the goal of all students becoming literate. We are not suggesting that the evaluation of teachers would not be a part of this movement; rather, evaluation under a Responsibility stance would begin with self-evaluation, leading to peer evaluation, and would find ultimate accountability in the determination of students and parents, and not in the actions of supervisors or managers.

Perhaps our efforts in the past—both ours personally and the profession's at large—to argue for change through policy reform at the top have failed following Ben Goldacre's (2009) caution that "you cannot reason people out of positions they didn't reason themselves into" (p. xii). We have come to believe that the audience we are writing for in this chapter is not the old guard who stand at the top of the accountability hierarchy and find their power within it, but the new generation of teachers who have been burdened by lives in classrooms dulled by standards and testing. Knowing the history as it is represented in this chapter may inform strategies for this new generation to move forward in unlocking its future into a transformative period of trust and excellence. Starting at the grassroots level, these dedicated professionals can inspire parents, colleagues, and administrators to adopt the spirit of responsibility in their own roles as the key to success. An appreciation for shared responsibility across communities of practice emerging out of these local interactions can become the hallmark for excellence. Perhaps, then, the policymakers who shape context and the publishers who craft tools for teaching will be convinced that the time is ripe for a new approach and a new era in literacy education. Jeanne Chall (1967) recognized this same tension when she asked which is better: a good teacher or a good method (e.g., phonics, look-say, or language experience)? She wrote: "Good teaching is always needed. But a good method in the hands of a good teacher—that is the ideal" (p. 308). We wonder, in the end, whether the question we posed in the title regarding trust in teachers or trust in programs may come to be answered: both! If it ever becomes necessary to make a forced choice between the two, as teachers, we know which side we will choose.

REFERENCES

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is—Or might be—The role of curriculum materials in teacher learning and instructional reform. *Educational Researcher, 25*(9), 6–8, 14.

Bereiter, C., & Engelmann, S. (1966). *Teaching disadvantaged children in the preschool*. Englewood Cliffs, NJ: Prentice-Hall.

Berman, P., & McLaughlin, M. W. (1976). Implementation of educational innovation. *Educational Forum, 40*(3), 345–370.

Bond, G. L., & Cuddy, M. (1962). *Good time for us, Companion first reader*. Chicago, IL: Lyons & Carnahan.

Bond, G. L., & Dykstra, R. (1967). *Coordinating Center for first-grade reading instruction programs* (Final Report of Project No. X-001, Contract No. OE-5-10-264). Minneapolis, MN: University of Minnesota.

Buenker, J. D., Burnham, J. C., & Crunden, R. M. (1977). *Progressivism*. Cambridge, MA: Schenkman.

Bynack, V. P. (1984). Noah Webster and the idea of a national culture: The pathologies of epistemology. *Journal of the History of Ideas 45*(1), 99–114.

California Learning Assessment System. (1994). *Elementary performance assessments: Integrated English-language arts illustrative material*. Sacramento, CA: California Department of Education.

Carnine, D., Silbert, J., Kame'enui, E., & Tarver, S. (2009). *Direct instruction reading* (5th ed.). New York, NY: Pearson.

Carver, R. P. (1990). *Reading rate: A review of research and theory*. San Diego, CA: Academic Press.

Chall, J. (1967). *Learning to read: The great debate*. New York, NY: McGraw-Hill.

Coleman, D., & Pimentel, S. (2012). *Revised publishers' criteria for the Common Core State Standards in English language arts and literacy, grades 3–12*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers. Available at www.corestandards.org/assets/Publishers_Criteria for_3-12.pdf

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

Cremin, L. A. (1957). *The republic and the school: Horace Mann on the education of free men*. New York, NY: Teachers College Press.

Davis, E. A., Palincsar, A. S., Arias, A. M., Schultz, A. Bismack, L. M., & Iwashyna, S. K. (2014). Designing educative curriculum materials: A theoretically and empirically driven process. *Harvard Educational Review, 84*(1), 24–52.

Edmonds, R., & Fredericksen J. R. (1978). *Search for effective schools: The identification and analysis of city schools that are instructionally effective for poor children*. Cambridge, MA: Harvard University, Center for Urban Studies.

Egbert, R. L. (1981). *Some thoughts about Follow Through thirteen years later*. Lincoln, NE: Nebraska University. (ERIC Document Reproduction Service No. ED244733)

Elmore, R. F. (1977). *Follow Through: Decision-making in a large-scale social experiment*. Unpublished doctoral dissertation, Harvard Graduate School of Education.

Elmore, R. F., & McLaughlin, M. W. (1988). *Steady work: Policy, practice and reform in American education*. Santa Monica, CA: Rand.

Filler, L. (1965). *Horace Mann on the crisis in education*. Yellow Springs, OH: Antioch Press.

Flanagan, M. A. (2007). *America reformed: Progressives and progressivisms, 1890s–1920s*. New York, NY: Oxford University Press.

Flesch, R. (1955). *Why Johnny can't read*. New York, NY: Harper and Row.

Goldacre, B. (2009). *Bad science*. London, UK: Fourth Estate.

Good, R. H., & Jefferson, G. (1998). *Contemporary perspectives on curriculum-based measurement validity*. New York. NY: Guilford Press.

Good, R. H., & Kaminski, R. A. (2002). *DIBELS oral reading fluency passages for first through third grades* (Technical Report, No. 10). Eugene, OR: University of Oregon.

Gould, L. L. (1974). *The Progressive era*. Syracuse, NY: Syracuse University Press.

Gray, W. S., Arbuthnot, M. H., et al. (1940–1948). *Basic readers: Curriculum foundation series*. Chicago, IL: Scott Foresman.

Gray, W. S., Arbuthnot, M. H., Artley, A. S., Monroe, M., et al. (1951–1957). *New basic readers: Curriculum foundation series*. Chicago, IL: Scott Foresman.

Hoffman, J. V., & Pearson, P. D. (2000). Reading teacher education in the next millennium: What your grandmother's teacher didn't know that your granddaughter's teacher should. *Reading Research Quarterly, 35*(1), 28–44.

Hoffman, J. V., & Rutherford, W. (1984). Effective reading programs: A critical review of outlier studies. *Reading Research Quarterly*, *20*(1), 79–92.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. F. (1978). No simple answer: Critique of the Follow Through evaluation. *Harvard Educational Review*, *48*, 128–160.

Improving America's Schools Act. (1994). Washington, DC: Government Printing Office. Available at www.ed.gov/legislation/ESEA/index.html

Kaestle, C. (1983). *Pillars of the republic: Common schools and American society, 1780–1860*. New York, NY: Hill and Wang.

Kendall, J. (2011). *The forgotten founding father: Noah Webster's obsession and the creation of an American culture*. New York, NY: G.P. Putnam's Sons.

Krajcik, J. S., & Davis, E. A. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, *34*(3), 3–14.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4–16.

Maccoby, E. E., & Zellner, M. (1970). *Experiments in primary education: Aspects of Project Follow-Through*. New York, NY: Harcourt Brace Jovanovich.

McLaughlin, M. W. (1976, February). Implementation as mutual adaptation: Change in classroom organization. *Teachers College Record*, 77, 339–351.

McLaughlin, M. W. (1990, December). The Rand Change Agent study revisited: Macro perspectives and micro realities. *Educational Researcher, 19*, 15.

Mendes, E. (2010). *U.S. Parents want teachers paid on quality, student outcomes. Gallop Politics*. Available at www.gallup.com/poll/142664/parents-teachers-paid-quality-student-outcomes.aspx

Monaghan, J. E. (2005). *Learning to read and write in colonial America*. Amherst, MA: University of Massachusetts Press, in association with the American Antiquarian Society.

National Commission on Excellence in Education. (1983). *A nation at risk: A report to the nation and the Secretary of Education, by the National Commission on Excellence in Education*. Washington, DC: Author.

National Governors Association (NGA) Center for Best Practices & Council of Chief State School Officers (CCSSO). (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors. Available at www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

National Research Council (NRC). (1999). Testing, teaching, and learning: A guide for states and school districts. Committee on Title I Testing and Assessment, Richard F. Elmore and Robert Rothman, editors. Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425. (2002). Available at www. ed.gov/offices/OESE/esea

Nichols, S. (2012). High-stakes testing: Does it increase achievement? *Journal of Applied School Psychology. 23*(2), 47–64.

Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.). New Haven, CT: Yale University Press.

Pearson, P. D. (2013). Research foundations of the Common Core State Standards in English language arts. In S. Neuman & L. Gambrell (Eds.), *Quality reading instruction in the age of Common Core State Standards* (pp. 237–262). Newark, DE: International Reading Association.

Pearson, P. D., & Hiebert, E. H. (2013). Understanding the Common Core State Standards. In L. Morrow, T. Shanahan, & K. K. Wixson (Eds.), *Teaching with the Common Core Standards for English Language Arts: What Educators Need to Know, Grades PreK–2* (pp. 1–21). New York, NY: Guilford Press.

Preston, J. (1989). Female aspiration and male ideology: School-teaching in nineteenth-century New England. In A. Angerman, G. Binnema, A. Keunen, V. Poels, & J. Zirkzee (Eds.), *Current issues in women's history* (pp. 171–182). New York, NY: Routledge.

Purkey, S. C., & Smith, M. (1983). Effective schools: A review. *The Elementary School Journal*, *83*(4), 426–452.

Resnick, L. (1993). as cited in J. O'Neil, On the New Standards Project: A conversation with Lauren Resnick and Warren Simmons. *Educational Leadership, 50*(5), 17–21.

Rice, J. M. (1893). *The public-school system of the United States*. New York, NY: Century.

Rice, J. M. (1913). *Scientific management in education*. New York, NY: Hinds, Noble & Eldredge.

Robinson, H. M., Monroe, M., Artley, A. S., & Greet, W. C. (1960–1962). *The new basic readers: Curriculum foundation series*. Chicago, IL: Scott Foresman.

Robinson, H. M., et al. (1967). *Open highways readers*. Chicago, IL: Scott Foresman.

Robinson, H. M., et al (1965). *Wide horizons readers*. Chicago, IL: Scott Foresman.

Rollins, R. (1980). *The long journey of Noah Webster.* Philadelphia, PA: University of Pennsylvania Press.

Rosenshine, B., & Furst, N. (1971). Research on teacher performance criteria. In B. O. Smith (Ed.), *Research in teacher education,* (pp. 37–72). Englewood Cliffs, NJ: Prentice Hall.

Ruggles, A. M. (1950). *The story of the McGuffeys*. New York, NY: American Book.

Smith, N. B. (1965). *American reading instruction*. Newark, DE: International Reading Association.

Stallings, J. (1976). How instructional processes relate to child outcomes in a national study of follow-through. *Journal of Teacher Education*, *27*(1), 43–47.

Stallings, J. A., & Kaskowitz, D. H. (1974). *Follow Through classroom observation evaluation (1972–1973)*. Menlo Park, CA: Stanford Research Institute.

Stebbins, L. B., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation: A planned variation model, Volume IV—A, An evaluation of Follow Through.* Cambridge, MA: Abt Associates.

Taylor, F. W. (1913). *Principles of scientific management*. New York, London: Harper & Brothers.

Tidwell, C. J. (1928). *State control of textbooks*. New York, NY: Columbia University Press.

Tulley, M. A., & Farr, R. (1985). Textbook adoption: Insight, impact and potential. *Book Research Quarterly*, *1*(2), 4–11.

Tyack, D., & Cuban (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.

U.S. Department of Education. (2009). *Race to the top: Executive summary.* Washington, DC. Available at www2.ed.gov/programs/racetothetop/executive-summary.pdf

Weber, G. (1971). *Inner city children can be taught to read: Four successful schools* (CBE Occasional Papers No. 18). Washington, DC: Council for Basic Education. (ERIC ED 057 125)

Webster, N. (1797). *The American spelling book: Containing an easy standard of pronunciation. Being the first part of a grammatical institute of the English language. In three parts* (19th ed.). New York, NY: Durham.

Webster, N. (1848). *An American dictionary of the English language.* Springfield, MA: George and Charles Mirriam.