PROJECTING FORWARD: LEARNINGS FROM EDUCATIONAL SYSTEMIC REFORM

Luther S. Williams
Margaret Cozzens

Printed in the United States

Any opinions, findings, conclusions and/or
recommendations herein are those of the authors and do
not necessarily reflect the views of the
National Science Foundation.

Published by

Comap, inc.
175 Middlesex Turnpike, Suite 3B
Bedford, MA 01730
781-862-7878
www.comap.com

ISBN-10: 0-9971490-4-3
ISBN-13: 978-0-9971490-4-3

Table of Contents

Preface

The National Science Foundation (NSF) initiated the design and provided the support for comprehensive complex system-wide reform of science, mathematics, and technology teaching and learning. The effort was achieved through four programs: the Statewide Systemic Initiatives (SSI), the Urban Systemic Initiatives (USI; focused on the 25 largest urban public school districts), the Local Systemic Initiatives, and the Rural Systemic Initiatives (RSI; consortia of rural school districts in a specific geographic domain), and, over time, the NSF Comprehensive Regional Centers for Minorities Program.

This volume begins with a detailed retrospective of the programs' goals, elements, accountability structures, and progress that were realized, while complying with a set of well-enumerated and uniformly employed programmatic anchors. These anchors — four process and two outcome "drivers" — were integral components of all systemic initiatives. Drawing on the findings of an array of external evaluation reports and case studies, the retrospective reviews the efficacy of systemic initiatives within the context of the specific program requirements and varied environments. It also gives a window into the contributions of specific variables to program achievements. In an effort to ascertain the scale and scope of the progression of systemic reform, issues such as the problem definition, elasticity of resources, programmatic redundancy, innovation transferability, design-redesign based on research and development on practice, and attention to gains in productivity are addressed.

The subsequent chapters discuss the outcomes and important lessons learned from these systemic programs, stressing documented student mathematics and science learning. These chapters are authored by experienced mathematics, science, and technology educators — all fully knowledgeable of the systemic efforts — and include former NSF systemic initiative program officers and division directors. These chapters also include such topics as reflections from inside NSF's programs of educational system reform, infusing equity in systemic reform, the Urban Systemic Initiatives in action, lessons learned and not learned from the systemic initiatives, the comprehensive regional centers for minorities program, technological innovation and urban systemic reform, the role of technical assistance, and how the game was changed. The last chapter provides a summary of the major findings and lessons learned from the systemic initiatives, with recommendations for the future.

Chapter 1

The National Science Foundation's Systemic Initiatives: A Retrospective Assessment

Luther S. Williams

National Context

Commencing in the late 1980s, there was an emerging understanding that the human resource development outputs of the nation's elementary and secondary education system (hereafter termed K–12) was greatly misaligned with powerful national and international trends, including the development of interdependent economies and novel workplace competencies for increasingly technologically oriented domestic and global work environments. By the latter portion of the twentieth century, these dramatic trends were mirrored by major societal changes, such as the end of the era of providing excellent K–12 education for only the privileged few! Rather, the knowledge and skills demanded by contemporary workplaces, taken in combination with citizenship-based activities requiring greater problem-solving capacity, demonstrable learning-transfer abilities, improved communication skills, information-gathering capacity, and improved knowledge-synthesis skills, obligated educational systems to afford a quality K–12 education for all students, particularly in mathematics and science. Yet, given this daunting challenge, K–12 systems possessing the aggregate capacities to deliver such an education constituted only the minority of the nation's school systems. That a less than ideal K–12 mathematics, science, and technology education had occasioned cumulative and quite negative impacts on the nation's scientific and technological enterprise is well summarized in *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*(National Academy of Sciences, 2007).

Systemic Programming

Under the visionary leadership of Director Erich Bloch, in 1989 the National Science Foundation (NSF) established the Directorate for Education and Human Resources (EHR). Among several other objectives, this new organizational structure was designed to promote the convergence of elements of science, mathematics, engineering, and technology education programs, and, as appropriate, forge consequential linkages between the previously described programs and the program activities designed to address underrepresentation. It was assumed that by creating a highly synergistic education and human resource development operation, shortages in human capacity would be addressed. Another effort, much informed by what was not accomplished through NSF funding during the prior decades, was designed to bring more dynamic dispositions, innovative strategies, and education research-based problem-solving approaches to EHR's program portfolio. Last, the feasibility of NSF pioneering systemic approaches, based on the theory of systemic reform of Smith and O'Day (1991), became the subject of careful and much extended discussions. To be sure, considerations of systemic approaches acknowledged the scale and complexity of the challenges of developing a comprehensive address of K–12 mathematics, science, and technology education in states, large urban school districts, and inter-sectoral rural school communities.

In particular, substantial attention was given to the daunting challenges that derived from the understanding that the multiple design elements and organizational constructs of systemic approaches would have their intended implementation in systems that more commonly operated within the confines of individualized categories of programming — for example, instructional materials, teacher preparation, informal science — and, more important, that rarely ever functioned at scale! Moreover, it was readily understood that many such school systems had standing practices that incentivized the conditions that produced the education deficit. The practices surrounding the expenditure of financial resources represented one such instance, which existed along with an array of other circumstances stubbornly at odds with a comprehensive K–12 mathematics, science, and technology education. However, the evident challenges notwithstanding, it was ultimately deemed appropriate that such a holistic, or all-systems, approach should be pursued. Such an approach would obligate the design and implementation of new and remarkably different K–12 mathematics, science, and technology education (teaching and learning) organizational models in, namely, systemic reform programs. Such a decision was reached with full knowledge that generally unproductive K–12 mathematics, science, and technology education operations had existed in the proposed funding sites for years and, therefore, the applicants to the proposed NSF systemic initiatives would be required to engage in creative and unexplored avenues of K–12 education with a relatively high potential for failure! In addition to the possibility of failure, the traditional NSF grantees, drawn from the higher education sector, were likely to expect and evaluate outcomes before they were warranted given the complexity of systemic reform projects. Nonetheless, the proposed undertakings were thought to represent an unusual opportunity to realize truly transforming or groundbreaking outcomes and, in this way, the systemic initiatives programs were initiated.

However, given the assumed risk, an accountability process was designed as an integral element of each systemic initiative award, the specifications of which were elaborated in each grant award instrument — a cooperative agreement (NSF,

EHR 1991; NSF, EHR 1994; NSF, EHR 1995. The use of a cooperative agreement in this instance might be viewed as the analog to the NSF Research Directorates funding; for example, the construction and operation of a major astronomical facility. This is in contrast to the regular research award mechanism used for a principal investigator grant addressing an authentic discovery process. Most of the requisite resources and/or implementation elements were potentially available for use by systemic initiative proposal developers and eventually awardees. These included mathematics and science standards-based curricula, sets of classroom-tested mathematics and science instructional materials, an array of elementary and secondary level teacher professional development models, the availability of state as well as national mathematics and science assessment systems — for example, National Assessment of Educational Progress (NAEP) — disaggregated to at least three grade levels within the K–12 teaching and learning continuum, and emerging sets of relevant learning technologies — some of demonstrable efficacy, an available cadre of program evaluators, and research reports/publications addressing various facets of mathematics, science, and technology education. Therefore, it was reasonable to expect that the results or outcomes occasioned by the progressive implementation of the systemic initiatives would lead to appropriate and different program adjustments, if reengineering were progressively undertaken. In a sense, there was no missing basic discovery process needed! As will be discussed, in systemic reasoning much needed to be learned. In further acknowledgment of the transformative exercise in which the potential awardees were being invited to engage, a planning grant was awarded to assist in the development of an effective implementation proposal (NSF, EHR 1991; NSF, EHR 1994; NSF, EHR 1995); and upon the award of a grant and the initiation of the proposed implementation processes, the NSF Systemic Reform program staff regularly conducted site visits.

As the assistant director of the NSF Systemic Reform program, I participated in such site visits to more than 12 urban school performance sites, as well as the states of Massachusetts and Louisiana, the Commonwealth of Puerto Rico, and schools of the Mississippi Delta and rural sectors of Appalachia. In addition, an "Urban Systemic Initiative (USI) Newsletter" was developed and used to disseminate information about program developments among and between the USI project leaders (PIs); USI superintendent conferences were held as were regional and national SSI, USI, and RSI conferences devoted to various facets of the overall implementation process — for example, data management, exploration of problem-solving of site or award-specific issues or differences in implementation approaches, local (site) evaluation, and the utility of applied research groups on-site.

Finally, through an award to an entity independent of each of the USI school districts, technical assistance was provided to the sites on a regular basis. In addition, the National Institute for Science Education, University of Wisconsin–Madison was funded to conduct research specifically undergirding systemic reform as such proceeded in the SSI, USI, and RSI performance sites; and resultant research monographs (Clune 1998, Kahle 1998, Webb and Weiss 2000) were disseminated to each of the programs. And, through funding by the EHR Directorate's Division of Research, Evaluation and Communication, several external organizations conducted comprehensive evaluations of each of the programs (Crasco et al. 2005, Kim et al. 2000, Kim et al. 2001, Kim et al. 2005, Malcolm et al. 2005, Yin et al. 2005, Zucker 1998, Zucker et al. 1998). Accordingly, at a minimum, the expectation of an answer to the following questions, yearly, was explicitly conveyed via the cooperative agreement: *(a) "Is the initiative anchored by the required continuum of standards-based curricula-instructional materials-instruction assessment? (b) Has the initiative served to promote the necessary programmatic synergy and productive participation/engagement/collaboration of all relevant stakeholders to yield a single coherently operated (unitary) system? and, (c) Is the initiative truly transforming as measured by significant positive impacts on the mathematics and science learning outcomes of all students?"*

Thus, the EHR Directorate of NSF, iteratively implemented a series of systemic initiative programs over the decade of the 1990s. The funding chronology was as follows: first, the SSI program focused on leveraging the K–12 science and mathematics education operation of an entire state and providing the necessary state infrastructure (e.g., in Connecticut, Ohio, and Louisiana); second, the USI program was designed to address systemic reform of K–12 mathematics, science, and technology education in the 25 largest urban school districts with the largest number of school-age children living in poverty (e.g., in Chicago, Detroit, and Miami-Dade County); and, third, the RSI program targeted large rural geographic domains (e.g., the Mississippi Delta, which comprises the rural school sectors of the states of Mississippi, Arkansas, and Louisiana).

Key Elements of the Systemic Reform Initiatives

While each NSF program solicitation was geared to the applicant domain (SSI, USI, and RSI), each was — at a minimum — expected to contain the following:

- The overall design should disallow fragmented or incomplete approaches. Rather, the program required a paradigm shift as it was expected to engage the total K–12 mathematics and science education enterprise.
- At the core of the proposal design was a curriculum-instruction-assessment system that was structured to be fully consistent with the K–12 mathematics and science standards of the National Council of Mathematics and National Research Council, respectively.
- Aligned and synergistic participation was required of all relevant entities (individual classrooms; schools; school systems and districts; faculty and administrators; community resource providers; local, regional and national government entities).
- Sustained convergence of human, financial, administrative, and governance resources was required for a productive systemic approach.

- On an annual basis, each participating systemic reform entity was expected to conduct rigorous examinations of outcomes related to its enumerated objectives with assessment metrics and employ the findings and accomplishments to make changes.

The outcomes reports were to be part of the required annual report, along with a "Program Effectiveness" exercise conducted by the EHR systemic reform program officers and external experts.

The intent of this annual exercise was to rigorously assess progress and formulate an improvement plan for the next fiscal year, thereby enabling a critically important iterative and continuous improvement process. In sum, through activities undertaken as described in the list above, mathematics, science, and technology education systemic reform awardees were expected to implement a comprehensive program designed to promoted fundamental, innovative and coordinated changes in financing, governance, management, mathematics, and science content to achieve equitable conduct of aggregate programming, resulting in demonstrable and sustained gains in mathematical and science learning by all students.

Monitoring Program Implementation

Neither the proscribed intentions of the EHR/NSF nor the quality of the program plans (proposals) submitted to NSF (and subsequently judged by external reviewers to be sufficiently meritorious to receive funding) necessarily ensured orderly implementation of the proposed work. In fact, the variance between that which was proposed and that which was actually implemented proved to be a major challenge in the instance of each of the systemic initiatives, the USIs in particular. This challenge occurred even though the preparation and submission of the systemic initiatives were preceded by multiple program-staff–conducted proposal development sessions and consultations related to specific elements of the solicitations and the terms and expectations that would be enumerated through the cooperative agreement. In fact, this was not surprising given the scope, scale, and uncommon design of the systemic initiatives, taken in combination with the more traditional allocation modes by which school districts or systems acquired program resources provided by various local and some federal funding entities. These findings notwithstanding, the systemic reform/EHR program staff elected not to reduce the intended dimensions of the systemic approaches given its commitment to answering the following question: Is systemic reform of K–12 mathematics, science, and technology education, as elaborated by EHR/NSF, a realistic undertaking in the designated performance sites? It is important to note that EHR remained steadfast in its commitment to test the feasibility of systemic reform even in the face of much pressure to break rank with the unitary system approach in favor of a less challenging piecemeal exercise!

Thus, in an effort to enhance the likelihood of a full-scale examination, the EHR/Division of Systemic Reform program staff eventually employed a two-tiered system to (a) initially monitor preliminary implementation of core systemic reform activities, and (b) subsequently conduct a year-end assessment of progress. The decision to implement this two-tiered process was very much influenced by the substantial variability in scale and scope of the systemic reform activities of the first and second SSI cohorts, funded in FY1991 and FY1992, respectively. Therefore, subsequent to the initial experience with cohort I of the SSIs, as well as the USIs (cohort I funded in FY1994), and RSIs (cohort I funded in FY1995), the cooperative agreements were revised and enumerated a specific set of deliverables or unit time; defined the performance threshold expected for continued support beyond the fiscal year in question; and firmly conveyed restrictions regarding the use of NSF funds for the support of activities incongruent with the goals and objectives of the systemic initiatives, as proposed, approved, and funded. As a consequence of employing this assistance and accountability process, several cohorts I and II SSIs that evidenced considerable disconnect with the programs' intents were orderly phased out, as were several USIs a few years later.

Initial Monitoring of Implementation of Core Systemic Reform Activities

Systemic Reform program staff assessments, performed after a few months of implementation, included the following: (a) evidence of standards-based mathematics and science curricula, teacher professional development, instructional program, and the correspondingly appropriate student assessments; (b) progress with the establishment of proposed partnerships; (c) sufficiency of human and financial resource leveraging; (d) plans for the use of technical assistance for schools or districts requiring such support; (e) requisite leadership development; (f) support for developing, sharing, or converging resources in a unitary teaching and learning system; (g) evidence of congruence of all the relevant policy and governance structures to enable the pursuit of common goals; and (h) demonstration of the feasibility of the trajectory of the scale-up process per domain (e.g., school district, governance unit, state, county, city, rural sector).

Critical Year-End Assessment of Progress

The annual Program Effectiveness reviews assessed (a) evidence (categorical and aggregate) that the program has occasioned measurable enhancements in student achievement; (b) evidence that the program both served and intended to serve with positive impact all students of the service domain (state, county, city, or rural sector); (c) documentation of implementation of a comprehensive, multigrade, standards-based curriculum; (d) the existence of a coherent set of policies, consistently employed, supporting the systemic reform educational system; (e) progress in the convergence of all resources in support of the systemic reform program through a focused and unitary strategy; (f) the existence of purposeful and sustained use of system indicators (quantitative and qualitative) for iterative assessment of progress as a total system serving all learners; and (g) evidence of broad-based support from all relevant segments of the community (domain) in question. Initially, the review process for annual program effectiveness revealed reports of elastic quality, and the initial implementation review process frequently resulted in the provision of specific guidance or assistance to sites in question. In 1996, in response to this significant finding, the Systemic Reform Division developed the "instrument for annual report of progress in systemic reform," which included six critical developments, termed "drivers"; four were "process developments" that enabled reform and two were "outcome drivers or developments" (NSF, EHR 1996). Succinctly, the process and outcome drivers are described below.

Process Drivers

1. Implementation of comprehensive, standards-based curricula as represented in instructional practice, including student assessment, in every classroom, laboratory, and other learning experiences provided through the system and its partners.

2. Development of a coherent, consistent set of policies that support provision of high quality mathematics and science education for each student; excellent preparation, continuing education, and support for each mathematics and science teacher (including elementary teachers); and other learning experiences provided through the system and its partners.

3. Convergence of the usage of all resources that are designed for or that reasonably could be used to support science and mathematics education — fiscal, intellectual, material, curricular, and extracurricular — into a focused and unitary program to constantly upgrade, renew, and improve the educational program in mathematics and science education for all students.

4. Broad-based support from parents, policy makers, institutions of higher education, business and industry, foundations and other segments of the community for the goals and collective value of the program, based on rich presentation of ideas that undergird the program, the evidence gathered about its successes and its failures, and critical discussions of its efforts.

Outcome Drivers

5. Accumulation of a broad and deep array of evidence that the program is enhancing student learning, as assessed through the use of a set of indices (e.g., achievement test scores, higher level course pass rates, advanced placement tests taken, college majors, portfolio assessment, research experiences, ratings by summer employers). In the specific instance of student test scores, awardees shall report, on an annual basis, the results of student mathematics and science achievements in a multigrade context for SI-impacted schools, districts, states, and interstate sectors, as appropriate, relative to appropriate cohort entities (non- SI districts, the state), all of which are defined by the performance baseline.

6. Improvement in the achievement of all students, including those historically underserved, as evidenced by progressive increments in student performance characterized by the requisite specificity of the SI as a catalytic resource and the appropriateness of the attendant attributions.

The annual SI reporting exercise focused on these six processes and outcome drivers, and it was supported by the development and use of a companion instrument, *Core Data Elements for Systemic Initiatives* (NSF, EHR 1998).

Results

While there were significant variations across the 25 states, the Commonwealth of Puerto Rico, 7 rural interstate sectors, and 22 large urban school districts, overall the systemic programs achieved substantial improvements in students' mathematics and science learning outcomes and the intended requisite teaching and learning infrastructure. In fact, the impacts of the programs were truly transformational in several states (Connecticut, Louisiana, Massachusetts, the Commonwealth of Puerto Rico), cities (Cincinnati, Dallas, El Paso, Miami-Dade, Memphis), and the Alaska Native and Mississippi Delta rural sectors. Each of these SI programs was characterized by the implementation of most or all components or elements of the systemic reform construct. Drawing on the results of external evaluation reports for the three SI programs, I will cite selected summary findings for each of the systemic initiatives, with major emphases on the USI, and then seek to connect them to major lessons learned.

Statewide Systemic Initiative (SSI) Program

In June 1998, SRI International submitted its five-year evaluation report of the SSI program to the Division of Research, Evaluation, and Communication, EHR/NSF (Zucker et al. 1998).The report said:

> The SSI Program has been a valuable test bed for the concept of standards-based systemic reform. As the theory of systemic reform suggested, each of the SSIs did, in fact, support a series of coordinated efforts affecting a number of different components of the education system. The most successful SSIs had very ambitious, comprehensive plans for systemic reform and were able to carry out those plans effectively with significant impacts. The least successful SSIs had designs that were narrowly focused and/or experienced implementation, quality control or management problems. Overall, the SSI program moved classroom practice in directions that are generally considered an improvement over past practices. After five years, the SSI program has left behind a legacy of new or improved curriculum frameworks, changes in a variety of state policies, new institutions and partnering arrangements, an increase in the number of competent state and local leaders of reform. . . . Overall, the SSI program has provided substantial contributions to standards-based, systemic reform of K–12 mathematics and science education.

In an April 1998 interview with Education Week, Andrew A. Zucker, a co-project director of the SSI Program Evaluation, stated, "For $50 million a year, it was a small investment in federal terms; for that money, the program garnered a very respectful, really impressive amount of program change" (Zucker 1998).

Among the SSIs, Connecticut, Massachusetts, the Commonwealth of Puerto Rico, and Louisiana were the most successful. Among the four SIs, Louisiana, a state with the nation's highest poverty rate (26 percent), the highest percentage of families headed by a single parent, the second-highest percentage of children living with parents without full-time, year round employment, and the sixth-highest percentage of births to mothers with less than 12 years of education, is an SSI exemplar (Davidson 2005). In a May 2004 Louisiana SSI program report, "Louisiana's Comparative Scores on the National Assessment of Educational Progress (NAEP), Louisiana's Improving Achievement in Mathematics," it is reported that for the 1992, 1996, 2000, and 2002 administrations, Louisiana students showed "the third best gain of 42 states at the 4th grade and second best gain of 41 states at the 8th grade" (Louisiana Systemic Initiative 2005). This noteworthy outcome is not surprising, as the Louisiana SSI operated for its entire tenure under the combined leadership of the PI, Dr. Kerry Davidson, a Louisiana Board of Regents professional, the Louisiana State Superintendent of Education, and the Louisiana Commissioner of Higher Education. In fact, there were only two Louisiana state superintendents during the life of this SSI. State Superintendent Arveson (1992–-96) stated the following, "I have been in education since 1942. In all those years, I have not seen a federal program as productive as LaSIP in its effects on children in the classroom" (Louisiana Systemic Initiative 2005LA Report, NAEP, 1992-2003). State Superintendent Picard (1996–) noted that "our recent reform efforts have their strategic foundation in the concepts of standards-based and systemic reform within a comprehensive school and accountability system. In part or whole, these concepts were originally conceived and advanced through the LaSIP" (Louisiana Systemic Initiative 2005). The Louisiana SSI scaled up from **391 schools, 783 teachers, and 55,000 students in 1992** to 850 schools, 3,700 teachers, and 240,000 students in 1996; then **to 1,300 schools, 7,000 teachers, and 400,000 students in both 2001 and 2003** (Davidson 2005). Discounting a large fraction of Louisiana rural schools that were only marginally impacted by LaSIP (these rural school districts were components of the Mississippi Delta RSI and the New Orleans Parish Public School district, which received a USI award), the 400,000 students represented about **75 percent of the total enrollment**. It is important to note that the rural school districts of the Mississippi Delta RSI, initiated several years after LaSIP, greatly benefited from the replication of field-tested LaSIP reform elements and activities. The minimum impact of LaSIP on the New Orleans school district was owed in part to its separate USI award; however, through a comprehensive site visit to both LaSIP and the New Orleans USI, I learned of the relative lack of a systemic reform engagement, which limited reform progress within the New Orleans school district. Equally important, it was evident that a powerful set of policy, governance, resource, and sociopolitical factors particularly served to disadvantage the New Orleans School District as compared to other Louisiana state school districts. This finding was determined not to be restricted to this single urban school district, USI site. With respect to convergence of financial resources (Davidson 2005) of the $46.9 million awarded to LaSIP through competitive grant programs between 1991 and 2002, $16 million was provided by NSF, $12.9 million was funded by the Louisiana Board of Elementary and Secondary Education, $11.5 million by the Louisiana Board of Regents, and $1.36 million by the Louisiana legislature. Not surprising given his history of working systemically, Dr. Davidson, in a 2005 presentation, proposed the continuation of LaSIP through a collaboration between it and Louisiana GEAR-UP, funded by a U.S. Department of Education grant (Davidson 2005).

Similarly, by 1998 other SSIs had evidenced considerable progress, including substantial and stable gains in student mathematics and science achievement, massive teacher professional development, intense focus on classroom practices, and other implementation strategies common to the SSIs (Corcoran et al. 1998, Shields et al. 1998, Zucker and Shields 1997, Zucker and Shields 1998). Regarding scale-up or students impacted, the **Puerto Rico SSI** served 600,000 students, or **92 percent of all student population**. The SSI principal investigator, Dr. Manuel Gomez, a senior University of Puerto Rico System academic officer, had forged a productive collaborative structure for the operations of the NSF funded EPSCoR and other federal grant programs across the university system. Thus, he effectively employed the university campuses, located in diverse settings throughout the Commonwealth of Puerto Rico, as regional resource support entities to the SSI school districts or individual schools, thereby greatly facilitating the scale-up process. In an effort to serve all students, the Puerto Rico SSI even dared to successfully extend its reach to a school district geographically removed from the Commonwealth mainland by some 40 or more miles. During a site visit, I observed students of African and Spanish origins situated in a decidedly third world school infrastructure reminiscent of the one-to-several rooms that Rosenwald Schools provided for African American students during the era of stringently enforced Jim Crow policies and practices in the Alabama Black Belt counties. Yet, under the instructional guidance of quite competent teachers (each having participated in extensive SSI- provided teacher professional development), the students were quite engaged in SSI-based algebra and middle school science teaching and learning!

The Massachusetts SSI provided standards-based, inquiry-centered mathematics and science education programs **for 70 percent of the total number students in the state**. In addition, Massachusetts student mathematics and science achievement increased at an impressive rate per assessment administration. In fact, if disaggregated from the U.S. averages, the state of Massachusetts would occupy placement among the exemplary performers in international assessments of student mathematics and science achievement, and still does today! However, in the context of SIs, such exemplary student achievement did not extend to the Boston Public Schools, a USI awardee.

In sum, Louisiana, Massachusetts, Commonwealth of Puerto Rico, and other successful SSIs were advantaged by the convergence of policies and resources (I would also include sociopolitical capacity) in the support of a singular or nearly singular enterprise. Specifically, for the period of 1991–98, the $283 million provided by NSF leveraged about $1.4 billion from the participating states, However, in all relevant SSIs, the least well-served domains within the state were the large urban school districts (note New Orleans and Boston, mentioned above), a matter that is further addressed in the Major Lessons Learned section.

Rural Systemic Initiative (RSI) Program

The RSI program, initiated in 1994, three years after the SSIs, sought to support systemic reform of mathematics and science education in rural areas in which more than 30 percent of school-age children lived in poverty. The findings that some SSIs had, by design, only modestly served rural schools and the acknowledged yet poorly understood near absence of substantive educational innovations in rural sectors were the principal factors leading to the establishment of the RSI program. Based on the results of comprehensive evaluations of the RSI program and detailed analyses of annual program effectiveness reports, each of the RSIs made reasonable progress with certain components of systemic reform and thus contributed to improvements of K–12 mathematics and science education in the designated rural regions of the nation. In particular, the implementation of standards-based curricula and inquiry-guided teacher professional development were generally quite impactful. Likewise, elements such as policies, partnerships, and building the necessary community infrastructure emerged as strong components of the RSIs.

Owing to a range of circumstances (exceedingly high rate of teacher and administrator turnover, the general lack of parental participation in school activities, the historic and cumulative consequences of geographical isolation, and much limited internal expertise), the RSIs achieved less than expected, especially with regard to progression to even 50 percent of the intended scale (Kim et al. 2005). Nonetheless, regarding capacity and expertise, two of the RSIs benefited from either productive linkage with a SSI (the Mississippi Delta RSI employed many of the systemic reform elements and strategies developed by the Louisiana and Arkansas SSIs) or programmatic linkages with another major initiative and state agencies (program collaboration between the Alaska RSI and the Annenberg Rural Challenge, and a formal agreement between the Alaska RSI and the Alaska Commissioner of Education for implementation of the initiative throughout rural schools). And, through the agreement with the Department of Education, mathematics and science education became a priority, with special attention given to Native issues. In fact, the Alaska RSI pioneered a demonstrably effective instructional approach, "infusing native ways of knowing," into the science, technology, engineering, and mathematics (STEM) curricula, which was adopted by other RSI sites that served indigenous students, including those of the Navajo Nation. An interesting and unexpected outcome of the RSIs that engaged Native Tribal Colleges was the transition of several of the colleges toward STEM-centric programs, a few of which now offer baccalaureate degree programs. Providing program coverage throughout the diverse and multistate service domains of several RSIs proved to be an understandable challenge. In an effort to address this challenge, the Appalachian RSI (ARSI), covering rural sectors of the states of

Kentucky, Virginia, West Virginia, North Carolina, and Ohio, made increased use of technology; implemented a six-state consolidated plan for utilizing state resources and categorical funds through the "Improving America's Schools Act" (U.S. Department of Education, 1994); operated the ARSI through the use of five regional, university-based resources collaboratives; and realized substantial progress over the duration of the RSI. Reform elements left on the ground in Appalachia include (a) a large pool of trained teacher partners for capacity-building; (b) university-anchored resources collaboratives; (c) distributed leadership teams tasked with data-driven program audits used as the basis for the elaboration of school improvement plans; (d) effectively leveraged and substantially converged resources; (e) the adoption and several-year use of appropriately revised policies and practices; (f) the multiyear use of substantive teacher professional development as the implementation nexus; (g) a standards-based mathematics and science education program; and (h) the benefit of considerable support by parents, community leaders, and business and industry (the retired CEO of Ashland, Inc., chaired the ARSI Advisory Council for the duration of the initiative), primarily occasioned by annual reports of improved student mathematics and science achievement. As specifically noted in one of the RSI program evaluation reports, "The Appalachian Collaborative Center for Learning, Assessment and Instruction in Mathematics contributed almost $1 million for professional development" (Kim et al. 2005). In addition, the convergence of the resources provided by the U.S. Department of Education–funded Appalachian Technology in Education Consortium and those of the Appalachian Regional Commission, a first of its kind of collaboration across rural school sectors of the participating states, served to powerfully enable the ARSI as the principal instrumentality for the promotion of mathematics and science education reform! Given its accomplishments and the on-the-ground traction, the ASRI-enabled work has continued through the awarding of two highly relevant projects: *The Appalachian Math and Science Partnership* and the *Leadership for Master Teachers.*

Overall, the RSI program can be accurately described as a considerably incomplete undertaking. Yet the findings evidencing enhanced instructional infrastructures or capacities, until now decidedly absent in most of the sites; substantial convergence of resources in support of a unitary agenda; positive and sustained improvements in student mathematics and science achievement; and the much increased high school completion rates by Native American students suggest that a disruption of the status quo by the audacious interventions represented by the RSI were not without merit!

Urban Systemic Initiatives (USI) Program

While the USI program was expected to benefit from the experience gained from the SSIs, its genesis was premised on the fact that **urban school systems enrolled approximately 50% of U.S. public school K–12 students**. Moreover, as noted in the USI solicitation, there is a well-documented "disparity between the academic performance of these students and that of their counterparts in suburban schools," an achievement gap that is not independent of "uneven allocation of resources," including experienced (and highly competent) teachers," as well as the paucity of "advanced courses, curriculum materials, instructional equipment and facility" (NSF, EHR 1994).

Accordingly, the program offered its participants a new paradigm for change; those who elect to participate (proposers/awardees) become partners in an effort to address the nation's need to make provision for all students. Reform involves the ability and willingness to make hard decisions (to break rank with history/tradition), take risks, resist simple solutions, and ensure that the development of children (specifically, in mathematics and science learning) is the central focus of the enterprise. Prior to submitting USI proposals, the USI applicants were advised to focus on a novel set of conceptual tools/models/strategies for addressing equity (Anderson et al. 1998); to provide access to high quality mathematics and science content; to ensure a competent instructional workforce; and to enable the transition from piecemeal or fractional approaches to sustained emphases on the strategic leverage points that promote systemic reform — all with the requisite understanding that systems are or will be complex, dynamic, continually adapting, even evidencing points of resistance that must be accommodated or nulled. Nonetheless, the ultimate and inelastic goal was a high performance unitary mathematics and science teaching and learning system serving all students!

The following summary is based on reports of several external evaluations of the USIs (Kim et al. 2001, Kim et al. 2005, Malcolm et al. 2005, Yin et al. 2005) and comprehensive analyses of 21 sites representing the three cohorts. From an initial base of serving **2.3 million students in academic year 1993–94, in academic year 1998–99, the number of students served by the 21 sites increased to 4.5 million; the attendant number of schools and teachers participating in the USI programs were 5,500 and 158,163, respectively** (Kim et al. 2001). Participation is defined as providing a high quality, standards-based curriculum delivered by teachers who had received significant professional development (i.e., 155,000 of the 250,000 teachers that comprised the instructional workforce of the USIs were provided 20–300 hours of professional development) in support of an inquiry-based approach to the teaching and learning of mathematics and science. Teacher quality was deemed critically important and an average of 49 percent of USI funds was spent on professional development activities. For example, for the 1998–99 school or academic year, 75 percent (118,399) of teachers working in USI schools participated in some form of professional development (Kim et al. 2001). That major progress was realized by the USI sites in providing teacher professional development is highlighted by the following examples. By year three, the Memphis USI had provided professional development for 4,000 teachers, 320 K–12 administrators, and better than 250 elementary, middle, and high school counselors; comparable levels of teacher professional development occurred in the San Diego, Detroit, San Antonio, and Phoenix USIs; the Fresno USI provided all 4,800 of its mathematics and science teachers with up to 120 hours of professional development. The Detroit and Dallas USIs served 100 percent of their students, 161,000 in 264 schools and 158,000 in 215 schools, respectively; and in the Miami-Dade USI, 100 percent of the elementary schools, 90 percent of the middle schools, and 60 percent of the high schools had implemented a teacher professional development–enabled standards-based curriculum within the respective classrooms. A comprehensive survey of the enacted curriculum of eight USI sites showed that classroom practices were revised to comport with the standards-based mathematics and science curricula and the learning acquired through the teacher professional development offerings (Kim et al. 2001).

Most of the school districts enacted policies essential to the orderly progression of the USIs. As examples, Memphis eliminated remedial high school science and mathematics courses; Detroit passed and implemented a K–12 resolution that required the teaching of the USI-centric mathematics and science courses at all grade levels, as well as coordination of professional development across all programs and aligned with curriculum changes; the three districts of the El Paso USI required three years of mathematics and science; and the nine districts of the San Antonio USI established common policies to support their collaborative partnership in support of the USI instructional program; state policies effecting preservice education were enacted that resulted in new certifications in mathematics and science in Dallas, Fresno, and Detroit; and Dallas instituted a policy change that increased the graduation requirement to include four years of science and mathematics. Regarding resource convergence, across the USIs, USDE categorical, foundation, business/industry, and USI district fiscal, infrastructure and human resources were synergistically employed to support the unitary USI program. The NSF USI investment leveraged over $1.1 billion between fiscal years 1996 and 1998,with $547 million for fiscal year 1999 alone (Kim et al. 2001). As an aggregate reform program, the USI realized SI reform elements, such as management of change, despite the frequent changes in superintendents and the massive collection, analyses, and effective utilization of data; implementation and stabilization of the standards-based curriculum-instruction-assessment combinations; and the establishment of partnerships and the development of internal reform capacity/expertise.

Urban reform in education included a plethora of sociopolitical and economic issues that often negatively impacted urban communities, including urban schools. Adding to this a national currency of positions or observations disdainful of the academic performance of urban K–12 students, it was exceedingly important that the USIs forge productive partnerships and that these partnerships provided broad-based support and necessary resources for convergence on a unitary mathematics and science education program, the USI. According to the USI evaluation reports,

> Every USI reported numerous partnerships. The infrastructure of partnerships built around each USI site provided strong support for systemic changes based on their unique capacity, resources and experiences. The partners and school districts interacted as part of a unified effort to

> promote and support large-scale education reform. These partnerships created avenues to connect all interested community members to the school system. The USI partnerships, organized as partner groups, included higher education over 130 universities and colleges, corporations, foundation research centers and laboratories (more than 125), informal institutions(in excess of 110 institutions), and parent and community organizations, e.g., Family Mathematics and Science programs. (Kim et al. 2001)

Examples of such partnerships include the Dallas USI higher education partners — 13 state and private institutions that provided assistance with professional development for preservice teachers, recertification of current teachers, and custom-designed coursework that meets the state and district curriculum framework. The Memphis USI included the University of Tennessee, Christian Brothers University, University of Memphis, and LeMoyne-Owen College and provided student summer enrichment activities in science, technology and algebra, tutorial programs, and Saturday Academies to support achievement in algebra, geometry, precalculus, physical science, biology, chemistry, and physics. The USIs forged linkages among disparate supporters and partners for the development and operation of community activities to support the initiatives; such partnerships existed with local institutions of higher education, botanical gardens, nature centers and museums, science and engineering industries, medical and health-care facilities, and federal research laboratories. The El Paso USI included the engagement of the University of Texas at El Paso as a principal resource. Texas Instruments was involved with 16 USIs, and IBM, Annenberg, and other private foundations provided financial support that increased from $2.5 to $60 million over five fiscal years.

Mathematics and science achievement increased across all districts, demonstrating the impact of the USI program directly on students. Over 75 percent of the USIs evidenced a positive correlation between student achievement and the length of time cohorts of schools participated in the USI program. Nearly all USIs reported increased student performance using state or national assessments; and improvements in student enrollment in and completion of higher level courses as low level courses were eliminated (Kim et al. 2001). Specifically, in one USI program evaluation report, it is stated that, for the 21 USI sites,

> noteworthy gains in student achievement [were realized], with the greatest gains seen in school districts that had participated in the USI program for the longest period of time. Urban students in the USI districts have substantially increased their enrollment rates in mathematics and science gate-keeping courses and higher-level courses. Underrepresented minority students made even greater enrollment gains than their peers during the same period, resulting in reduced enrollment disparities. Assessment test results show that the USI students made gains in science and

> mathematics achievement, while reducing achievement gaps among racial/ethnic groups. The increasing numbers of 11th and 12th grade students taking the college entrance examinations (AP, SAT, ACT) indicate that more students have aspirations of pursuing post-secondary education. (Kim et al. 2001)

Thus, we can conclude that the exemplary sites constitute proof of what is possible in urban school districts when they choose to implement a coherent, highly focused, unitary, and intentional mathematics and science teaching and learning system, enabled and supported by all the program drivers. Equally noteworthy are the substantial increases in mathematics and science learning outcomes as a result of these USI programs, which served all or nearly all students.

Comprehensive Partnerships for Mathematics and Science Achievement (CPMSA) Program

The CPMSA program, initiated one year in advance of the USIs, was funded by the Human Resource Development Division of EHR to improve mathematics and science education in medium-sized cities, with a focus on improving learning outcomes for underrepresented students. As the program progressed, it transitioned to a more systemic approach, inclusive of standards-based curriculum, instruction, and assessment and the objective of improved achievement by all students. From 1993 through 1998, the program grew to 25 CPMSAs (e.g., Omaha, Newport News, Akron, Kansas City, Laredo, Birmingham) with a **total enrollment of 927,607 students**, **and over 29,000 teachers in more than 1,300 schools**. As reported in the external evaluation report (270, the program experienced "significant gains in mathematics and science learning outcomes (as measured by improved achievements on district/state assessments), increments in the rates of enrollment in and completion of advanced mathematics and science courses, progressive increases in advanced placement course enrollments, and improves performance on the SAT and ACT examinations" (Kim and Crasco, 2003, pp. 1–3). See Chapter 6 for reflections on the CPMSAs.

Local Systemic Change (LSC) Program

The LSC program was initiated in 1995 by the Elementary and Secondary Education Division of EHR. With the core objective of substantially improving instruction in science and mathematics through teacher professional development, the LSC program differed from the division's prior teacher professional development programs by emphasizing preparing teachers to implement specified exemplary mathematics and science instructional materials in the classrooms of entire schools and school districts. Accordingly, the LSC program provided all teachers of a school or district professional development for a minimum of 130 hours of professional development over the duration of the project. As with the SI programs, the LSC programs were to align policy and practices and included activities that undergirded and/or supported reform, including a shared vision, substantive and iterative self-studies of implementation, participation by the requisite partners, and participation in an exhaustive evaluation conducted by an external reviewer.

From an initial cohort funded in 1995, the LSC program grew to a total of 88 projects funded by 2002. According to the evaluation report (Banilower et al. 2006, pp. 85–86), the LSC program had significant impact in several areas, including high quality and sustained professional development; teachers and classroom practices, classroom adoption and use of exemplary instructional materials, and progress in building capacity and infrastructure needed to support and, more importantly, sustain reform.

Major Lessons Learned

Before indicating specific lessons learned, it is important to emphasize some of the features of the context within which the SI programs were designed and implemented. First, the SIs (SSI, USI, and RSI) obligated a systemic approach. Second, the dimensionality of the undertaking was substantially heightened by the a priori enumeration of a common set of essential, challenging systemic drivers. Third, although many SI performers were devoid of the required system dynamics, each was provided substantive assistance and guidance and held accountable to implement a unitary coherent program. If they did not or could not, they were phased out in an orderly way. Fourth, the complexity of the SI enterprise was daunting for the sites, often due to the pressure of implementing multiple components simultaneously, requiring them to define and promote the relevant education policy changes; leverage resources and build meaningful partnerships; and ensure convergence of all relevant resources to the support of one system. In addition, they had to explicate and disseminate best practices; promote the dissemination of "early phase" systemic reform school, district, or state-level models; and conduct regular data collection and analyses. Finally, they were expected to employ the emerging mathematics, science, and technology research base. Lessons were learned in this context. What is needed is standards-based curriculum-instruction-assessment and systemic thinking.

Standards-Based Curriculum-Instruction-Assessment

Setting clear and ambitious mathematics and science standards was shown to be an imperative to sustained improvements in student learning outcomes and the orderly implementation (in the classrooms) of proscribed critical systemic components. However, viewing the SIs as an aggregate, there was a lack of uniform emphasis on the need for the students to build deep knowledge of mathematics and science content. Enabling the students to develop a good understanding of mathematical and scientific habits of mind, without a substantial content knowledge base on which to apply such insights, proved to be less effective than intended. Also, the results of a range of assessments affirmed the criticality of ensuring that all policies were aligned with and served to integrate diverse educational transactions, such as teacher preparation, teacher professional development, general and content pedagogy, use and utility of learning technologies, student assessment systems, and equity in all dimensions.

Systemic Thinking

Systemic reform was shown to require competent, sustained, committed, and accomplished leadership/management of (a) forces to support the reform and to counter other forces that tried to dislodge the unitary enterprise, (b) powerful holders of the policy development and implementation processes who might seek to retain their power, (c) moneyed interests positioned obverse to the core SI agenda, (d) those resistant to change, and (e) an array of other nonenabling circumstances present in a system characterized by complexity.

In addressing the complexity of systemic reform, knowledgeable, committed, and sustained system leadership was shown to be crucial. For example, one of the USI evaluation studies found a positive correlation between the school district superintendent's understanding and sustained support of the USI project and continuity of employment and the "rates of student participation in mathematics and science courses" of high quality (Kim et al. 2001). This is quite a reasonable finding since the school district leader ultimately must (a) propose and secure school-board approval of polices bearing mathematics and science requirements, (b) convey essentiality of systemic thinking and actions to existing and potential external program partners and funders, (c) productively converge human and financial resources in support of a unitary mathematics and science teaching and learning enterprise, (d) enable the district mathematics and science administrator(s) and principals, and (e) value and support the requisite instructional workforce. Thus, not surprising, it was reported that in analyses of several variables, "The effect of a stable and supportive superintendent is twice as large as the combined effects of the USI project director's reporting relationship with the superintendent and the size of the USI staff " (Kim et al. 2001). Moreover, such exemplary SSI, USI, and RSI leaders procured the participation of local higher education mathematics and science expertise, regularly sought technical assistance afforded to the sites through an EHR-contracted SI technical assistance provider, and engaged the entire system of professionals in an effort to heighten their understanding of systemic thinking, inclusive of innovation transfer, productivity, accountability, and sustainability of the system.

All Students

The transition from serving a select few to providing a highly competent education for all students requires a firmly executed policy structure that enables equitable resourcing of each participating school with all drivers of the SSI program. The inclusion of this matter as a mandatory one owes to the essentiality of an equitable provision of education and, importantly, the fact that its omission from SIs would have changed the program's goals and objectives from systemic reform to a support mechanism designed to promote improvements of selected units or fractions of a system.

Reform Capable Infrastructure

Mobilizing the expertise needed to assist the various sites in achieving success was challenging; however, the sites were able to generate incremental improvements at the beginning and subsequently cement the gains and proceed to the next phases of the reform process. Moreover, the sites employed the required annual report of student achievement results to motivate further changes. The annual Program Effectiveness exercise's overall results were used as a vehicle to deeply probe the status of most, if not all, systemic elements. Thereby, this aggregate process allowed the sites to iteratively assess the status of the reform infrastructure and then undertake a redesign in preparation for the next implementation period (fiscal year). Viewed with the benefit of hindsight, NSF might have been well served to have procured the professional services of "system dynamics" experts to conduct system progression analyses in each site. In that regard, it would seem useful to employ the core tenets of Cohen and Ball, as referenced in Chapter 7, of a dynamic and complex system transitioning to scale. Such a study would seem to be critically important for "every school, every student" in a system that itself was comprised of an aggregate of solutions!

Closing

I close this reflection (retrospective) with a paraphrase of a historic utterance attributed to Sir William Wallace in fourteenth-century Scotland, which I (privately held) upon the initiation of the USI program, in particular. For we had to "take the sterling bridge of doubt, racism, and neglect and (by demonstrable performance) convince all detractors, on-lookers and even seekers of the truth" that the USI was not the next in a long series of education programs in urban school districts. Rather, "it sought to unlock the key to the nirvana of the student's survival." Today I posit that by initiating and insisting upon the substantive and comprehensive conduct of such challenging systemic reform programs, there has emerged a set of useful experiences, findings, and knowledge bases; and these outcomes might properly serve as an important guide to future actions undertaken to substantially improve students' mathematics and science learning in comparable teaching and learning domains.

References

Anderson, B.; Campbell, P. B.; George, Y.; Jolly, E.; Butler, J.; Kreinberg, N.; Lopez-Ferrao, J.; and Taylor, G. (1998, reprinted 2000). Infusing Equity in Systemic Reform: An Implementation Scheme. National Science Foundation.

Banilower, E. R.; Boyd, S. E.; Pasley, J. D.; and Weiss, I. R. (2006). "Lessons from a Decade of Mathematics and Science Reform: A Capstone Report for the Local Systemic Change through Teacher Enhancement Initiative." Horizon Research, Inc.

Clune, W. H. (1998). "Toward a Theory of Systemic Reform: The Case of Nine NSF Statewide Systemic Initiatives." National Institute for Science Education, Research Monograph No.16.

Clune, W. H.; Millar, S. B.; Raizen, S. A.; Webb, N. L.; Bowcock, D. C.; Britton, E. D.; Gunter, R. L.; and Mesquita, R. (1997). "Research on Systemic Reform: What Have We Learned? What Do We Need to Know?" National Institute for Science Education, Research Monograph No. 4.

Confrey, J. (2017). "Technological Innovation and Urban Systemic Reform: –Designing for Change," chap. 7.

Corcoran, T. B.; Shields, P. M.; and Zucker, A. A. (1998). A Report on the Evaluation of the NSF's Statewide Systemic Initiatives (SSI) Program, the SSIs and Professional Development for Teachers. SRI International.

Crasco, L. M.; Kim, J.J.; and Leavitt, D. J. (2005). "Rural Systemic Initiatives: Key Indicator Report 2004." Systemic Research, Inc.

Davidson, K. (2005). "Louisiana Systemic Initiative." Presentation to the LIGO Education Advisory Committee.

Kahle, J. B. (1998). "Reaching Equity in Systemic Reform: How Do We Assess Progress and Problems." National Institute for Science Education, Research Monograph No. 9.

Kim, J. J., and Crasco, L. M. (2003). "Overcoming Challenges in Urban Education, Comprehensive Partnerships for Mathematics and Science Achievement." Systemic Research, Inc.

Kim, J. J.; Crasco, L. M.; Blank, R. K.; and Smithson, J. (2000). "Survey Results of Urban School Practices in Mathematics and Science, 2000 Report: An Evaluative Study of NSF's Urban Systemic Initiatives." Systemic Research, Inc.

Kim, J. J.; Crasco, L. M.; Bride, A.; Leavitt, D. J.; and Weiner, J. K. (2005). "Rural Systemic Initiatives: Bringing Science and Mathematics Excellence to Rural Classrooms, RSI Program: Highlights and Case Stories of Seven Sites." Systemic Research, Inc.

Kim, J. J.; Crasco, L. M.; Smith, R. B.; Johnson, G.; Karantonis, A.; and Leavitt, D. J. (2001). "Academic Excellence for All Urban Students: Their Accomplishments in Science and Mathematics, USI Evaluative Study." Systemic Research, Inc. Louisiana Systemic Initiative. (2005). "*Louisiana's Comparative Scores on the National Assessment of Educational Progress, 1992–2003*."

Malcolm, S. M.; Abdallah, J.; Chubin, D. E.; and Grogan, K. A. (2005). "System of Solutions: Every School, Every Student." American Association for the Advancement of Science. National Research Council, Committee of Science, Engineering and Public Policy (USA). (2006). *Rising Above the Gathering Storm: Energizing and Employing America for a Bright Future*. National Science Foundation, Directorate for Education and Human Resources. (1991). *Statewide Systemic Initiative Program Solicitation.*

———. (1994). "Urban Systemic Initiative Program in Science, Mathematics and Technology Education."

———. (1995). "Rural Systemic Initiative in Science, Mathematics and Technology Education."

———. (1996). "Instrument for Annual Report of Progress in Systemic Reform."

———. (1998). "Core Data Elements Summary for Systemic Initiatives."

Shields, P, M.; Marsh, J.; and Adelman, N. E. (1998). "The SSIs' Impacts on Classroom Practice." SRI International.

Smith, M. S., and O'Day, J. A. (1991). "Systemic School Reform." In S. Fuhrman and B. Malen (eds.), Politics of Curriculum and Testing. Bristol, PA: Falmer Press.

U.S. Department of Education. (1994). "Improving America's School Act of 1994."

Webb, N. L., and Weiss, I. R. (2000). "Annual Report of the Study of the Impact of Statewide Systemic Initiatives." National Institute for Science Education.

Yin, R. K.; Davis, D.; Schmidt,; and Leavitt, D. J. (2005). "Final Report: Strategies and Trends in Urban Education Reform, Cross-Site Evaluation of the Urban Systemic Initiative Program."

Zucker, A. A. (1998). "Federal Funds 'Well Spent' on Science Effort, Study Finds." Education Week.

Zucker, A. A., and Shields, P. M. (1998a). "SSI Case Studies, Cohort I: Connecticut, Delaware, Louisiana, and Montana." SRI International.

———. (1998b). "SSI Case Studies, Cohort III, Arkansas and New York." SRI International.

Zucker, A. A.; Shields, P. M.; and Adelman, N. E. (197). "Student Achievement in the SSIs." SRI International.

Zucker, A. A.; Shields, P. M.; Adelman, N. E.; Corcoran, T. E.; and Goertz, M. E. (1998). "A Report on the Evaluation of the National Science Foundation's Statewide Systemic Initiatives Program." SRI International.

Luther S. Williams is distinguished professor of biology emeritus at Tuskegee University, at which he iteratively held positions as professor and director of the Integrative Bioscience PhD Program, dean of Graduate Studies, provost and vice president for Academic Affairs, and executive vice president. For the decade of the 1900s, he served as the assistant director of the Education and Human Resources Directorate, National Science Foundation, during the Statewide, Urban, and Rural Systemic Initiatives, the Local Systemic Change, and the Comprehensive the Regional Centers for Minorities programs development and implementation. Concurrently, he served as vice chair of the Federal Coordinating Council for Science, Engineering and Technology Committee on Education and Human Resources, White House Office of Science and Technology Policy. He subsequently served as the William Kemper Director of Education and Interpretation at the Missouri Botanical Garden, at which he launched the Garden Education Compact, a science and mathematics education collaboration between the Botanical Garden and a group of St. Louis City elementary and middle schools. Prior to Tuskegee, his academic positions included professor and president, Atlanta University; professor and Vice President for Academic Affairs, University of Colorado; professor and dean of the Graduate School of Arts and Sciences at Washington University, St. Louis; associate professor, Massachusetts Institute of Technology; and assistant, associate, and full professor and assistant provost, Purdue University. He received the PhD in microbiology/molecular biology from Purdue University and was a postdoctoral fellow at the State University of New York at Stony Brook.

Chapter 2

Systemic Reform: Vantage, Reflections, and a Prediction

Eric Hamilton

This chapter begins with observations about the history of the educational system reform movement underwritten by the National Science Foundation (NSF), primarily in the 1990s. It situates system reform as partial successor to well-known science, technology, engineering, and mathematics (STEM) programs for teachers and students for which NSF became especially well-known beginning in the 1960s. While not displacing those programs, system reform became an intellectual epicenter for agency investments in K–12 STEM education, up to the time of eventual transition to NSF's participation in the No Child Left Behind (NCLB) Act.

My overall thesis is that system reform was a tremendous and positive jolt to U.S. education; that its successes arose from a coherent and sound understanding of educational systems; and that its shortcomings, in retrospect, were relatively predictable and arose, as much as anything, from the impossibility of attempting to impart and hasten curriculum and instructional philosophies into practice before systems fully assimilated those philosophies. The system reform movement, before it was eclipsed by NCLB policy, offers powerful lessons and will, in any case, be retold in the eventual story of U.S. educational success. If it is not retold, it is because we will not have succeeded in our aspirations. Success will go through the many lessons of system reform.

The second part of this chapter addresses topics germane to looking forward in the context of the compelling lessons of system reform. The first topic is the so-called contemporary Finnish miracle — an example of an educational system that reformed itself and become a top international performer. Finland offers object lessons that are counterintuitive across ideological spectra for U.S. education reformers. These lessons fit the system reform movement closely but with important gaps to note. The second topic is the role of new findings about learning and new technologies to enhance learning, and the third is the pervasive and unmistakable role of assessment in formal education systems.

Each contributor to this book brings a distinct perspective. This chapter represents an admixture, as Dr. Luther Williams would say, of multiple professional roles and dispositions that are useful to disclose. I served as director of one of NSF's comprehensive regional centers for minorities, in Chicago, as the agency transitioned from an intervention-centric to a system-centric strategy for addressing achievement gaps and workforce shortages. I then served as the initial director of the Chicago Systemic Initiative before coming to NSF to serve as a program director and then as a division director for the research and evaluation division of the Education and Human Resource directorate that Dr. Williams led.

NSF shifted away from the systemic initiative strategy and into the Mathematics and Science Partnerships, or MSPs. I coauthored the initial Research, Evaluation, and Technical Assistance (RETA) "Dear Colleague Letter" associated with the MSP program. I currently run a series of research projects under NSF support, focusing on imagination and creativity in STEM teacher development and in student-teacher collaboration, and the Science Across Virtual Institutes (SAVI), which is an international collaboration network connecting learning scientists and learning technologists in the United States and Finland (http://innovations forlearning.net). The SAVI is aimed in large part at bringing lessons from Finland to the United States. Current work also focuses on education reform in low-income countries, with an increasing emphasis on system change, versus intervention strategies in low-income countries.

This context creates a potentially useful vantage point, of observing the transition into the system reform strategy and participation in it as a grantee, program director, and then a division director for Education and Human Resources (EHR) as the agency transitioned away from the system reform approach. The ascendancies of digital tools, social media, reconceptualizations of classrooms and other learning environments, and emergence of analyses related to successful systems internationally provide additional perspective and useful hindsight in crystallizing the trove of lessons from the system reform movement.

Current trends in U.S. STEM education are varied, are complex, and resist generalization. It is clear, though, that our education systems continue to underperform and to underserve, especially in economically marginal areas and in predominantly African American, Latino, and Native American populations. It is also clear that the welfare of students and of their future has become even more politicized than in the past. Considerations about political advantage — that is, considerations other than learning — dominate discourse and decision making. This is unworthy of our democratic ideals. Our national discourse dynamics are often toxic, derisive, and poorly informed. Marshalling sufficient will, resources, and knowledge to reverse the nation's tolerance of educational mediocrity will require solutions and pathways that I argue below have not yet been formulated or, at least, that have not been operationalized. But those solutions will encompass and adapt the many lessons of the systemic reform movement, or else they will reinvent them at great cost of time, treasure, and national well-being.

The Simple and Profound Insight of Systemic Reform

The systemic reform movement that NSF spurred a generation ago brought large-scale changes in participating local education agencies (LEAs) and state education agencies (SEAs), changes that altered life and learning for millions of K–12 students. The great insight of the movement was simple but profoundly deep: To change the outcome of systems, systems needed to change. Otherwise phrased, perhaps, this maxim was the turning point for NSF, the shift from interventions to systems. System reform enlarged the brief of NSF's involvement in educational enterprises significantly; it gave the agency leverage because its funding, small compared to the budgets of the systems it sought to alter, was discretionary and not obligated in service of maintaining the system. The disadvantage, as related elsewhere in this book, is that grant expenditures rarely translated into sustained system expenditures once the external funding concluded.

An effort combining the expertise of the various Educational System Reform (ESR) team members, led by Dr. Williams, and revised continually in interactions with system grantees and those to whom NSF reports, resulted in the conceptual structure of drivers of system change. The formalization of the drivers was a significant watershed for all of us serving as ESR program officers, leading to an almost palpable sigh of relief to have in hand an analytic grid that encompassed, in near totality, the complex endeavors they were overseeing. The drivers provided a framework for stimulating system change by identifying and simplifying the management and assessment of multiple levers of change concurrently. It was adopted as a logic model in education reform literature (e.g., Clune 1998), but research directly addressing the driver construct, outside NSF-contracted evaluations, was thin, at least on the surface. Research addressing components of system reform, though, routinely highlighted the need for the very system alignments foundational to NSF's guidance and oversight of initiatives. Elkind (2004), for example, reviewing shortcomings of curriculum change in major reform efforts, concluded that misalignment between teaching and curriculum was the primary deficiency in those efforts. Similarly, Blumenfeld et al. (2000) cited misalignment or, more accurately, the inability of a system to sustain practical logistics and the day-to-day particulars of curriculum and technology shifts, as underlying systemic reform's most salient challenges and shortcomings. Misalignment of system drivers eventually became the defining element of UNESCO's assessment grid for progress in education reform in low-income countries (UNESCO 2013). The underlying theory of ESR was that coordinated progress in each system driver would result in lasting system change. This sound model did not have the political and social backing for the multi-decade effort that eventual resounding success would have required. Indeed, this was the sort of time line of national resolve and will that has led to success in high-performing systems such as Finland's, discussed below.

It is difficult to convey how intellectually powerful the shift to a systems model became for pursuing national interests for large-scale improvements in STEM fields. It certainly created an economic appeal — "scaling up" improvement so that life and learning for students is better as a matter of the system being itself rather than as a matter of special interruptions (interventions) to the system is undeniably alluring. There was a larger force field, though. Systems thinking organized around simultaneously reaching high achievement and eliminating achievement gaps animated profound moral, spiritual, and social sensibilities of countless educators, parents, policy makers, researchers, and others. The level of energy poured into systemic initiatives by those carrying out systemic reform exceeded by a wide margin the actual monetary compensation or salaries they received. There is no doubt that the availability of substantial financial resources drew many in, but the energy brought to bear when an agency devoted to the nation's scientific well-being promoted the notion that "all children can learn" fueled an enormous movement. That it did so in an intellectually coherent, morally clear voice, and made learning the prerequisite benchmark for winning resources, added to the excitement and determination of those engaged in the ESR movement.

I mention this intangible energy factor because, on first approximation, there was a deep resonance across the awardee community that is important to acknowledge and understand in any treatment of the system initiative history. And, on further review, finding and channeling that energy is essential to future endeavors.

While systems thinking was giving shape to NSF's investment strategies in education, it was also taking hold in other areas (e.g., Haines et al. 2004). In particular, the recognition of how human conceptual systems evolve, especially in science and mathematics, gave rise to sizable mathematics, science, and engineering education research communities in the area of conceptual models (Lesh and Harel 2003, Lesh and Kelly 1997, Roschelle and Jackiw 2000, Schorr and Lesh 2003). The science of complex adaptive social systems also made great strides in the 1990s, and toward the end of the decade important connections between the science of complexity and educational systems reform were at least being explored or were the subject of coherent conjecture (e.g., McElroy 2000, Goldstone and Sakamoto 2003). I highlight two of them here. The first is the notion that processes constituting a system yield higher level effects as a system than they would individually. System behavior has properties that transcend and differ from the behavior of system elements (Hamilton 2015, Sabelli 2006). This notion of emergent properties, of course, is a primitive axiom of systems but, like other primitive axioms, may have a more pervasive effect than first appears. System design may reflect multiple positive intentions — including educating learners — but, as systems evolve and take form, they develop inexorable propensities for self-preservation that can and do subordinate original intention to status secondary to self-preservation. Complex and powerful systems adapt to external threats to those systems. One observer of system reform wryly noted that "systems resist systemic reform systemically."

Systemic Drivers and Program Effectiveness Reviews

By the late 1990s, educational systems funded by ESR reported significant progress in each of the six system drivers. The Program Effectiveness Review, or PER, became the primary vehicle for detailing progress. The PER process included both a written report and a reverse site visit, by which the systemic initiative team came to NSF to give a 90-minute presentation. In addition, NSF published a "USI Summary Update" providing details of the progress of each system toward the goals of their respective initiative. The individual PERs and the Summary Update all reported important system progress. They may have been the only summative accountability devices that crystallized progress both in system drivers and in student achievement, even though the NSF investment was less than 0.1 percent of the budget of any of the systems in which an award was made. They also gave impetus to greater transparency in accountability and deeper reflection within systems on the need for and nature of learner-centered curriculum and instructional practices.

The second of two themes emerging from the study of complex social systems is that system jumps or alterations, when they do occur, are generally self-organizing. NSF set out to effect system change by which the initiative in each local or state education authority (LEA or SEA, respectively), or rural region, was the central organizing principle for change. It would be inaccurate to argue that these were top-down reforms. Every systemic initiative (SI) had a unique lineage, but there was an inherent tension between establishing a central organizing force for system alteration and establishing one that emerged from the properties of the system itself. The contrast is seen vividly in the light of so-called viral processes or media, but such a contrast is more accurately metaphorical than real — NSF did not seek to effect viral change in clicks or attentional cycles but rather to effect lasting system change. But viral phenomena are a useful way to think of the rapid accelerant nature of some changes in social norms, attention, or attitudes.

The analytic process NSF required of systems ostensibly involved breaking down the levers that control enterprise dynamics. The SI model evidenced analytic competence and could marshal stakeholders and resource alignments because it was internally coherent and represented best known sub-solutions and fell within a deeply held moral and social force-field of so many educators, researchers, and policy makers. It was incandescent.

Teachers, Curricula, and Assessment

But hindsight gives us some advantages not available at the time of implementation. That very incandescence and certitude that characterized the SIs may have isolated the SI movement from attention to realities and vulnerabilities that were inescapable in the 1990s. One of the most critical was the lack of externally accepted or validated conceptual frameworks and instrumentation to assess and measure large-scale efforts in which complex systems were the units of change

(Supovitz et al. 2000). Attribution of effect to cause was the most intractable challenge facing the ESR program staff.

In addition to this deficiency, the infrastructure to support reform curricula underlying system change was inadequate. At the time of SI implementations, NSF was quite aware of — and supporting — the need for teacher professional development to support curriculum transitions, but the agency did not have the resources, the charge, or the methodologies to effect large-scale changes that culturally profound paradigm shifts and related professional development would require. The systems had to effect those changes. Altering curriculum and instructional practices across the spectrum of everyday skilled and effective teachers to be the most versatile and accomplished is daunting. It requires ways of thinking about both learning and learners that differ significantly from the patterns most teachers have acquired over a lifetime. Teaching is a complex profession saturated with tacit cues and understandings driven by those life patterns. This is not to imply that teachers could not or did not shift, or that many were not already practicing learner-centric styles, but culture shifts and the development of the collective professional expertise needed before measuring success of the initiatives were not well understood. In a sense, the SI movement demanded that educators more deeply appreciate the process of student learning, that giving lectures and exercises (teacher actions), for example, do not intrinsically translate to learning (a student phenomenon). It also demanded reliability and validity in assessing long-term effects of professional development efforts; conceptual frameworks for doing so remain under-theorized and underdeveloped (Desimone 2009). This is not a technical or minor issue. Complexities in impact measurement strike at the heart of any effort to validate professional development reform and are amplified in system reform. Deeply internalizing a professional shift to more student-centered ways of teaching is multilayered and constitutes more than adopting a new curriculum and attending professional development workshops. Changing the name of the commercially produced curriculum obtained on a purchase order or adopting the rhetoric of a forward-looking instructional ethos is not the same as changing how teachers carry out the curriculum, and carrying out professional development programs is not the same as teachers developing professionally in a manner that leaves them able to realize a curriculum developer's vision and interpret it meaningfully in their own respective context. It is difficult to overestimate the impact of promised reform falling short because teachers were put into the position of implementing curricula and different instructional practices without owning and having thorough command of those curricula and practices. The recent National Academies' report, *Science Teachers' Learning: Enhancing Opportunities, Creating Supportive Contexts* (2016) offers a fine-grained study of how a vision for effective science teaching in the United States, for example, falls short of the reality. The study furnishes an important analytic grid for understanding the complex landscape or ecosystem of the current science teaching workforce, as a prerequisite to developing a series of articulated policy and practice proposals for developing and improving that workforce. During the system reform movement, this study's level of insight and detail was simply not yet available to inform each initiative's respective professional development strategy. Simply put, what could be called the science of professional development in science education was not as far along as it has come and certainly not as far along as it needed to be.

The challenge in systemic reform relative to professional development for curriculum implementation was mirrored on the assessment side. Assessment systems were — and remain — wholly inadequate for providing indicators for the breadth of competencies we expect students to have when they leave their schools. The tradeoff in student assessment between efficiency (easy-to-read scores, limited cost and intrusion of time for educators, and profits for publishers) versus effectiveness (capturing multidimensional and varied learning) almost always comes down in favor of efficiency, at the expense of more authentic assessment of student progress. Assessment for the purpose of identifying and accelerating learning progress and assessment for the purposes of so-called accountability are not intrinsically oppositional, but neither do they fully overlap nor are they especially compatible. One toxic aspect of discourse in U.S. education is that assessment writ large has become a proxy for issues that have nothing to do with education by dint of conflating these differing purposes of assessment.

Every systemic initiative was greeted with fanfare and celebration in the host LEA or SEA. High expectations for change were ubiquitous; the axiom that all children can learn was repeated endlessly and believed fervently by grantees and applicants for SI support. In retrospect, I do not believe we had the tools and knowledge to realize that vision nearly as fully as we had hoped. Our education system writ large is so powerfully structured to support a single lane of learning that even a seemingly radical disruption, such as removing academic tracking, is inadequate. Our production-style approaches still require learners to wrap themselves around the system, rather than structuring a system that wraps itself around the learner. We are only now seeing some of those possibilities, including awareness that well-executed personalized learning trajectories, enabled in part by new technologies and approaches, will routinely trump production-style formal structures to education (Hamilton and Jago 2010).

A Comparison Between Finland and the United States

Finland is often viewed as a haven for successful educational practice (Sahlberg and Hargreaves 2011). An extensive literature on how or why Finland's educational systems outperform peer systems often cites factors such as higher social prestige for teachers and less emphasis on testing. Still others focus on the relative homogeneity of Finnish students. Finland is a nation accustomed, in recent years, both to the cash inflow and intrusion of visitors curious to know how and why a country that did not purpose to become the top performing educational system became just that.

It turns out that some of the most sacred dogmas of educational reformers in the United States hold little sway in Finland. Finnish teachers do not earn more pay, by and large, than their U.S. counterparts. Class sizes are generally not smaller than in the United States. Per pupil expenditures in Finland are substantially lower than in the United States (Wolff et al. 2014). Students are not expected to begin their education in preschool; Finnish youngsters do not even start formal schooling until age 7. Each of these factors is resource-consequential and, in each case, the more successful system (Finland) expends fewer resources than the less successful system (United States). It is true that although Finland has its share of struggles in the full education of indigenous versus European populations, those struggles pale in comparison to the demographic challenges facing the United States. But issues of class size, teacher pay, per pupil expenditure, and start of schooling are worth noting. They are not determinative in system success in Finland. Finland leads the United States in system outcomes, but it lags behind in these subordinate component indicators.

Another factor often cited in comparing Finland to the United States is that Finnish students do not undergo the mandatory standardized and high stakes testing of their U.S. counterparts. While this is true, the difference is less determinative than derivative. That is, the difference reflects determinative underlying realities but is not crucial in and of itself. U.S. deficits relative to peer prosperous nations predated the testing regimens that came with the No Child Left Behind Act.

A more determinative set of factors is as simple and deep as the insight that to change system outcomes we must change systems. The teaching profession functions differently in Finland than it does in the United States. Of the many challenges education research faces in accumulating and interpreting findings and building strong warrants for results in one venue to inform policy in other venues, one finding remains consistent: the quality of teaching is the largest single factor in educational success. Teacher quality encompasses so many social and cultural dimensions, feedback loops, and self-modifying and self-reinforcing factors that it is futile to believe that unidimensional approaches to professional development or admissions standards or credentialing can produce a significant impact in and of themselves. They are necessary but woefully insufficient. A more holistic understanding of factors related to teaching must become pervasive. Finnish society, while not compensating teachers with average salaries higher than their U.S. counterparts, positions the profession as a much more valued national resource than does U.S. society. Respect matters. And professions that build respect tend to perform differently — and more effectively. More prestigious professions elicit higher performance from those predisposed to go into those fields and they attract the interest of high performers who would not otherwise consider them. To become a teacher in the United States, one pays often mortgage-level sums for a college education that is free in Finland, all to enter a relatively low-status profession in the United States that differs sharply from the higher status accorded to Finnish teachers. While there is a national curriculum in Finland, teachers there are accorded autonomy, latitude, and discretion in organizing their classrooms to achieve the curriculum, a form of professional freedom and respect rarely found in the United States (Sahlberg 2012), where teachers find themselves guilty until proven innocent on high-stakes testing and where the policy and political ambience of distrust undermines prospects for success.

All of these factors place teachers and teaching in untenable positions and polarize stakeholders whom the public should instead reasonably expect to work together to find solutions to complex educational challenges and learner needs. The fiscal deficits education systems face are considered the great challenge of the day, but it is easy to argue that the cash value of those deficits pale in comparison to the cash value of collaborative problem-solving capacity and operationalized trust that is squandered by the adversarial nature of our national educational enterprise. It is a bleak picture.

From a different perspective, though, the teaching profession has always benefited from the overwhelmingly positive sense of satisfaction its members experience in their day-to-day efforts (Lytle 1980, Nash and Ducharme 1983), what Dan Lortie originally coined as the "psychic reward" of teaching (cited by Ryan and Cooper 2012). The advent of new approaches, both conceptual and digital, is expanding the landscape of potential satisfaction for teachers, and it is in this landscape that the potential for disrupting systemic difficulties may be strongest. It turns out that some of the factors underlying the status profile of teaching in Finland may find different kinds of expression by teachers in the United States.

Summary

The chapters in this book highlight many of the lights by which our educational enterprise is unchanged, or has moved backward, in the years since the systemic reform movement. A common rhetorical device in conference papers or speeches about the pace of change in education is to show images of classrooms from the first part of the twentieth century and to suggest that not much has changed in the past century. But the device has become worn because education is changing more rapidly than it has in the past. The changes are multilayered. One transition: for the first time in history, the generation in schooling often surpasses their teachers in facility with the tools of knowledge sharing and storage (Hamilton 2015). This is a signal shift of the past 20 years. But other changes place teachers in more advantageous positions than they have ever enjoyed. Depending on the layer of analysis, K–12 education is at once stagnant and roiling with change. New tools are available to help teachers — learning management systems, a derided administrative facet of teaching, comprise an important shift in classroom practice. More important, innovations to enhance student learning, such as new cybertools, serious games, digital approaches to coursework, and slow but inexorable progress in understandings about learning, are leading to significantly different teaching practice. Emphases on teacher creativity, creativity as a critical literacy for learning, and teacher problem-solving will also contribute to the sense of agency and autonomy that is crucial to the profession (e.g., Mishra and Henriksen 2012, Hamilton 2013, Peppler and Bender 2013).

These alterations are not going to bring success to our educational system, nor will they address intractable social and political conditions. They may, however, help contribute a catalytic and disruptive effect. I predict that there will not be a significant improvement in rendering the U.S. K–12 educational enterprise a success without operationalizing the convergence of moral forcefulness, energy, resources, and deft use of system levers that characterized the systemic reform movement. But, as the expression goes, you can't steal first base. The reform environment in the United States is stuck, in ways that differ from the 1990s. To name just two, fault lines in education policy have become more politicized and challenging, and the contemporary era's digital reshaping of K–12 education practice is far more pervasive than it was in the 1990s when what was referred to as educational computing technology was relatively nascent. But even those dramatic shifts do not eclipse the most compelling conclusions or lessons from the system reform movement. I predict that innovations in learning sciences and learning technologies will alter the sense and experience of teacher professionalism, and the ascendance of these will be catalytic precursors to more sustained and systems-oriented changes in education. Without either this catalytic effect or the systems paradigm, our country will squander unprecedented opportunities for progressing to a more humane and successful education enterprise.

References

Blumenfeld, P., B. J. Fishman, J. Krajcik, R. W. Marx, and E. Soloway. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded project-based science in urban schools. Lawrence Erlbaum Associates. 35: 149–64.

Clune, W. (1998). Toward a theory of systemic reform: The case of nine NSF Statewide Systemic Initiatives. Madison: University of Wisconsin-Madison.

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. Educational Researcher 38(3): 181–99.

Elkind, D. (2004). The problem with constructivism. The Educational Forum, Taylor & Francis Online https://eric.ed.gov/?id=EJ773733

Goldstone, R. L., and Y. Sakamoto. (2003). The transfer of abstract principles governing complex adaptive systems. Cognitive Psychology 46(4): 414–66.

Haines, S., G. Aller-Stead, and J. McKinlay. (2004). Enterprise-wide change: Superior results through systems thinking. John Wiley & Sons.

Hamilton, E. (2013). Finding creativity and flow in a high-stakes assessment context. Irish Educational Studies 32(1): 109–17.

Hamilton, E. (2015). Advancing a complex systems approach to personalized learning communities: Bandwidth, sightlines, and teacher generativity. Journal of Interactive Learning Research.

Hamilton, E., and M. Jago. (2010). Toward a theory of personalized learning communities. In Designs for learning environments of the future, P. Rieman & M. Jacobson (eds.). Springer, pp. 263–81.

Lesh, R., and G. Harel. (2003). Problem solving, modeling, and local conceptual development. Mathematical Thinking & Learning 5(2–3): 157–89.

Lesh, R., and A. Kelly. (1997). Teachers evolving conceptions of one-to-one tutoring: A three-tiered teaching experiment. Journal for Research in Mathematics Education 28(4): 398–430.

Lytle, J. H. (1980). An untimely (but significant) experiment in teacher motivation. Phi Delta Kappan: 700–702.

McElroy, M. W. (2000). Integrating complexity theory, knowledge management and organizational learning. Journal of Knowledge Management 4(3): 195–203.

Mishra, P., and D. Henriksen. (2012). Rethinking technology & creativity in the 21st century: On being in-disciplined. TechTrends 56(6): 18–21.

Nash, R. J., and E. R. Ducharme. (1983). The paucity of the investment metaphor and other misunderstandings. Journal of Teacher Education 34(1): 33–36.

National Academies Teacher Advisory Council. (2016). Science teachers' learning: Enhancing opportunities, creating supportive contexts. National Academies Press.

Peppler, K., and S. Bender. (2013). Maker Movement spreads innovation one project at a time: Lessons Learned from the grassroots spreading of the "Maker Movement" can help us reimagine schools and foster a mindset of creativity and innovation in educational settings. Phi Delta Kappan 95(3): 2.

Roschelle, J., and N. Jackiw. (2000). Technology design as educational research: Interweaving imagination, inquiry & impact. In Research Design in Mathematics & Science Education, A. Kelly & R. Lesh (eds.). Mahwah, NJ: Lawrence Erlbaum Associates, pp. 777–97.

Ryan, K., and J. Cooper (2012). Those who can, teach. Cengage Learning.

Sabelli, N. H. (2006). Complexity, technology, science, and education. Journal of the Learning Sciences 15(1): 5–9.

Sahlberg, P. (2012). A model lesson: Finland shows us what equal opportunity looks like. American Educator 36(1): 20.

Sahlberg, P., and A. Hargreaves. (2011). Finnish lessons: What can the world learn from educational change in Finland? Teachers College Press.

Schorr, R. Y., and R. Lesh (2003). A modeling approach to providing teacher development. Beyond constructivism: A model and modeling perspective on teaching, learning, and problem solving in mathematics education, R. L. H. D. (eds.). Mahwah, NJ, Lawrence Erlbaum Associates, pp. 141–57.

Supovitz, J. A., D. P. Mayer, and J. B. Kahle. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. Educational Policy 14(3): 331–56.

Wolff, E. N., W. J. Baumol, and A. N. Saini. (2014). A comparative analysis of education costs and outcomes: The United States vs. other OECD countries. Economics of Education Review 39: 1–21.

Eric Hamilton received a PhD in mathematics education from Northwestern University and undergraduate and master's degrees from the University of Chicago. Originally tenured in computer science and mathematics at Loyola University Chicago, he led several award-winning efforts to increase the participation of underrepresented minorities in science and engineering professions before taking leadership of the Chicago Urban Systemic Initiative. After one year in that role, he moved to the National Science Foundation (NSF), where he oversaw urban and statewide systemic initiatives nationally before taking leadership of the Division for Research, Evaluation, and Communication. Dr. Hamilton subsequently initiated a research center in learning and teaching at the U.S. Air Force Academy and has managed numerous education and learning technology research projects inside and outside the United States over the past 10 years, at the Air Force Academy and at Pepperdine University, where he is a faculty member. He leads a network of 15 clubs in Kenya, Finland, Namibia, and the United States; the clubs collaborate in digital makerspace projects, as part of an NSF-funded research effort in virtual networks in science and engineering. He currently is a 2015–17 Fulbright Research Scholar with assignment in the Republic of Namibia.

Chapter 3

Infusing Equity in Systemic Reform: Influencing the Understanding of Equity and Excellence

Patricia B. Campbell and Eric Jolly

> Social inequities, that are inherent in today's educational systems, must be eliminated. Equity must be at the center of every reform effort so that it cannot be marginalized or isolated from learning experiences that lead to high achievement in mathematics and science by all students.
>
> —Anderson et al. 1998

This call to action —the elimination of social inequities—continues to resound. It grew out of frustration with the little progress that was being made in eliminating achievement gaps and of great hopes for the impact that the National Science Foundation's systemic initiatives (SIs), particularly the Urban Systemic Initiatives (USIs), would have on our nation's children. The call grew out of a 1996 National Science Foundation (NSF) Equity Planning Retreat, which brought together epresentatives of national organizations and SIs to create the framework for a document to assist leaders of the SIs and of other educational reforms in their efforts to pursue the goal of equity and excellence in mathematics and science education for all students. The resulting document, *Infusing Equity in Systemic Reform: An Implementation Scheme* (Anderson et al. 1998) was disseminated widely and 17 years later continues to be cited and is now available from amazon.com.

Infusing Equity in Systemic Reform presented a profoundly new approach to increasing equity and diversity. Its underlying premise, that a school, classroom, program, or curriculum was high quality only if the vast majority of students succeeded, was a major departure from earlier ideas. It grew out of the analogy that since cars that only ran for some drivers or trains that only took some passengers to their destinations would not be defined as high quality, then neither should educational programs that don't serve all of their students be defined as high quality. Student success, the success of girls and boys from all races and ethnicities, was put at the center of the definition of high quality.

Infusing Equity in Systemic Reform proposed three major changes in the ways that equity and outcomes were defined and implemented:

- Equity is not an add-on to a program; it is an essential and inherent component of high quality mathematics and science at all levels of education. Thus, it must be infused into every aspect of systemic reform (e.g., curriculum, professional development, assessment, policies, resources, partnerships, data gathering, evaluation).
- Reducing achievement differences across groups is valuable, but the longer-term goal is to serve all students so that outcomes can no longer be predicted based on sex, ethnicity, class, disability, or language group. If group membership predicts individual participation and achievement, then the system will be defined as inequitable and, therefore, the education provided will not be meeting the standards established by the state or school district.
- Outcomes are what matters. Process measures help us understand why we see or do not see improvement in a subject, but outcome measures tell us what students know and are able to do and who those students are.

These changes were proposed within a context in which NSF was "firm in its conviction that mathematics and science programs must be equitable and excellent in order to be of high quality. Unless all students are achieving in standards-based mathematics and sciences, no claim for high-quality education can be made" (Anderson et al. 1998, p. 6). However, 17 years later, "No claim for high-quality education can be made." The powerful ideas presented in *Infusing Equity in Systemic Reform* have had an impact; although, as seen below, when they were misunderstood or when there was only partial implementation, the impacts have not always been as positive as they could be.

Disaggregation of Data

One major change in education since *Infusing Equity in Systemic Reform* was published has been in the disaggregation of data. When achievement and other data were not broken out by such areas as sex, race/ethnicity, and disability, if disparities existed, it was not known because there were no data. Today achievement and other outcome data are routinely reported by basic demographic categories, including race/ethnicity, sex, and, in some cases, disability. Increasingly, these data are reported, so one can look at the interaction of race/ethnicity and sex. This allows one to look at, for example, changes over time for African American boys, white boys, African American girls, and other subgroups. This breaking down of data has been invaluable in showing where the problems are and the progress, or lack of progress, being made over time.

These data show where the problems are but not necessarily what the problems are or what can be done to resolve problems or inequities. To do this, data collection and analysis needs to include information on variables that can be changed, such as curriculum, pedagogy, and disciplinary policies. That was one of the recommendations of the 2011 NSF *Framework for Evaluating Impacts of Broadening Participation Projects (Framework)*, written by Beatriz Chu Clewell, one of the coauthors of *Infusing Equity in Systemic Reform*:

> NSF should expand its systematic and objective evaluation to assess, understand, and report the effectiveness and impact of its programs and policies on broadening participation by: Continuing to obtain, refine, and disaggregate data and factors related to the participation and advancement of persons from underrepresented groups in STEM education and careers. (Clewell 2012, p. 29)

As shown in the following example, disaggregating data when looking at outcomes and at factors that impact outcomes can lead to change:

> Pam Hammen got tired of staring at the data. The data were telling Hammen, the principal of Verona Area High School in suburban Madison, Wisconsin, that 55 percent of white students in her school were proficient in reading in tenth grade, as opposed to 13 percent of black students. At the same time, students of color were being suspended at much higher rates than their white peers. "It was obvious that we had an achievement problem, coupled with disproportionality in discipline," Hammen said. "We were not happy about that, and we set out to change." (School of Education News 2014)

Elimination of Disproportionality

Disproportionality, according to the Equity Project at Indiana University,

> may be defined as the over- or under-representation of a group in a category that exceeds our expectations for that group, or differs substantially from the representation of others in that category. Although concerns have historically tended to focus on issues of over-representation in special education or school suspension and expulsion, groups may also be under-represented in a category or setting (e.g., under-representation in general education settings or gifted education). (Equity Institute)

Infusing Equity in Systemic Reform introduced the role of disproportionality in achievement when it stated that "if students' achievement can be predicted by their gender, race/ethnicity or physical ability, then the system is inequitable and must be changed." Clewell's listing of metrics in the *Framework* reinforced this, explaining that for the research and education communities served by NSF, this would appear to imply, minimally, metrics related to the following: "absolute and relative (to the general population and to relevant availability pools) rates of participation in STEM research and education activities and professions by students, staff, faculty, and administrators drawn from underrepresented populations" (Clewell 2011, p. 30).

In our discussions for the *Framework* about when success is reached in diversity efforts, Campbell, Thomas, and Stoll explained:

> Defining success as "reducing the gaps while all gain" puts change in context but does not provide an end point — a point when success is achieved. One definition of success that includes an "end point" is parity, which is defined by Merriam Webster (2008) as "the quality or state of being equal or equivalent." In the cases of the first two areas, having access to the benefits of STEM knowledge and having access to STEM knowledge, parity could be with the population of the United States. That is, members of designated subgroups (i.e., underrepresented minorities [URM], women, people from rural areas) have the same access as does the population as a whole. For [studying STEM], parity may be the percentage of the subgroup studying STEM that reflects the subgroup's percentage in the general population. That is, if African Americans are 12.8 percent of the population, then parity in the third area [studying STEM] is African Americans being 12.8 percent of those studying STEM. (2011, p. 55)

Today, however, for the most part, education remains focused on "reducing the gaps while all gain" rather than eliminating disproportionality even though, in 1998*, Infusing Equity in Systemic Reform*, made it clear that looking at gaps was not enough: "If achievement is not proportional across identifiable subgroups of students, the selected curriculum is not meeting the desired goal of high-quality science and mathematics education for all students. If that is the finding, then other evidence about the curriculum must be sought. And the result may be to select another curriculum" (Campbell et al. 2011, p. 28).

While one can celebrate that the gaps are declining, albeit slowly, when looking at the amount of time it would take to reach parity in math and science achievement and participation there is little to celebrate.

Focusing on Outcomes

Infusing Equity in Systemic Reform stated unequivocally that "outcomes are what counts." In our follow-up piece, "Moving into the Mainstream," Campbell and Kreinberg (1998) took it a step further, recommending moving from "access and treatment at the core of accountability" to "outcomes at the core of accountability." Through efforts such as No Child Left Behind, that has happened. Outcomes have become the core of accountability.

This focus on outcomes is important, but equally important is what leads to those outcomes. *Infusing Equity in Systemic Reform* pointed out, "In systemic reform, it is important to understand that both student and school inputs affect student outcomes. Therefore, evidence for the presence of high-quality science and mathematics courses and for access and retention in them must precede any attempt to seek evidence for changes in student outcomes" (p. 28). This is, or should be, obvious.

Access and treatment are not outcomes; they are strategies to achieve outcomes. If courses are not available, students will not take them. If there are no laboratories, students will not be able to acquire laboratory skills. If teachers do not know the content, students will not either. Similarly, if students are ignored or not given adequate attention, they are unlikely to learn. If students are discouraged from continuing on in a field of study, they will unlikely do so. If financial or scheduling barriers exist, some students will fail to overcome these barriers. As Campbell, Thomas, and Stoll point out in the *Framework,* "Generally, student level [achievement] indicators can be identified across four dimensions, including: (a) participation, (b) retention, persistence, and success, (c) experiences, and (d) attitudes" (2011, p. 57).

And, of course, what the outcomes are is key. If outcomes are limited to standardized test scores and don't include the full range of skills, knowledge, and attitudes needed to continue on in STEM, then they are incomplete and will not necessarily lead to a qualified and diverse STEM workforce and a STEM knowledgeable public.

Equity as an Essential and Inherent Component of High Quality Mathematics and Science

Clewell and Campbell (2007) studied typical and highly effective elementary schools serving low-income, minority students in two urban areas. We found that equity was an essential component of the highly effective schools and that students in predominately, minority low-income schools could achieve at levels comparable to students in predominately white schools with few low-income students. In the highly effective schools, teachers as well as principals took responsibility for all students in the school, often calling them "our children." Teachers saw student learning as their responsibility as teachers and if the students weren't learning, it was the teachers' fault and the teachers' responsibility to figure out a way to improve student learning. Their schools' disciplinary policies were clear and consistently applied and their principals were instructional leaders. They were clear about the outcomes they wanted, and their expectations for the students were high. Their schools had educational philosophies, not necessarily the same philosophies but each had a clear philosophy and all were on board with it. The highly effective schools infused equity in their efforts and it worked.

In Closing

Equity was at the core of the systemic initiative effort. Through that effort, the interrelationship between equity and high quality STEM education was expanded and reinforced. In the years since then much has changed for both the better and the worse; however, the line in the sand that was drawn in 1998 in *Infusing Equity in Systemic Reform* remains the standard to which we need to aspire. “Unless all students are achieving in standards-based mathematics and science, no claim for high-quality education can be made.”

References

Anderson, Bernice; Campbell, Patricia B.; George,Yolanda; Jolly, Eric; Kahle, Jane Butler;

Kreinberg, Nancy; Lopez-Ferrao, Julio; & Taylor, Gwen. (1998, reprinted 2000). Infusing Equity in Systemic Reform: An Implementation Scheme. Arlington, VA: National Science Foundation.

Campbell, Patricia B.; & Kreinberg, Nancy. (1998). “Moving into the Mainstream: From Equity a Separate Concept to High Quality Includes All.” Washington, DC: AAAS.

Campbell, Patricia B.; Thomas, Veronica; & Stoll, Adam. (2011). “Outcomes and Indicators Related to Broadening Participation.” In Framework for Evaluating Impacts of Broadening Participation Projects, Beatriz Chu Clewell & Norman Fortenberry (eds.). Arlington, VA: National Science Foundation, pp. 54–63.

Clewell, Beatriz Chu. (2011). “Measuring Success and Effectiveness of NSF’s Broadening Participation Programs: Suggested Monitoring Metrics and Evaluation Indicators.” In Framework for Evaluating Impacts of Broadening Participation Projects, Beatriz Chu Clewell & Norman Fortenberry (eds.). Arlington, VA: National Science Foundation, pp. 42–53.

Clewell, Beatriz Chu, & Campbell, Patricia B., with Perlman, Lesley. (2007). Good Schools in Poor Neighborhoods: Defying Demographics, Achieving Success. Washington, DC: Urban Institute Press.

The Equity Project at Indiana University. Understanding Equity. (n.d.). http://www.indiana.edu/~equity/undequ.php.

School of Education News (June 6, 2014). MSAN Helps Districts Examine Ways to Tackle Racial Disparities in School Discipline. University of Wisconsin-Madison

http://news.education.wisc.edu/news/2014/06/06/msan-helps-districts-examine-ways-to-tackle-racial-disparities-in-school-discipline.

Skiba, R. J., & Rausch, M. K. (2006). Zero Tolerance, Suspension, and Expulsion: Questions of Equity and Effectiveness. http://www.indiana.edu/~equity/docs/Zero_Tolerance_Effectiveness.pdf.

Patricia B. Campbell, president of Campbell-Kibler Associates, Inc., has been involved in STEM educational research and evaluation with a focus on issues of race, gender, and disability since the mid-1970s. Formerly an associate professor of research, measurement, and statistics at Georgia State University, she has authored more than 120 publications, including coauthoring www.BeyondRigor.org; The Effect of Context on Student Engagement in Engineering; Seeking Effective Math and Science Education and Good Schools in Poor Neighborhoods; and Defying Demographics, Achieving Success and Engagement, Capacity and Continuity: A Trilogy for Student Success. An Association for Women in Science (AWIS) Fellow, Dr. Campbell received the Betty Vetter Research Award from Women and Engineering Program Advocates (WEPAN) and the Willystine Goodsell Award from the American Educational Research Association.

Eric Jolly is president and CEO of Minnesota Philanthropy Partners. Formerly the president of the Science Museum of Minnesota, he has published numerous articles, books, and curricula, and he has lectured around the world about the importance of STEM (science, technology, engineering, and mathematics) education in contemporary societies. Dr. Jolly works with a number of groups, including the National Science Foundation's Advisory Committee on Education and Human Resources, the National Research Council's Out-of-School Time STEM Learning Committee, the National Academy of Engineering's Committee on Implementation of K–12 Engineering Education in the United States, the North Star STEM Alliance, and the American Association for the Advancement of Science's Committee on Science and Technology Engagement with the Public. He is a member of the Institute of Museum and Library Services board, appointed by President Barack Obama.

Chapter 4

Thinking About the Urban Systemic Initiatives

Daniel Burke

The Urban Systemic Initiatives (USI) award program arose from the research-based findings that all children can learn and achieve at a high level and from the ethical conviction or imperative that children deserve an educational system that will enable them to do so. The National Science Foundation chose the use of systemic reform as the model required to implement the USI awards to achieve the necessary educational system to meet the students' needs (Darling-Hammond, Lieberman, & Miller 1992; O'Day & Smith 1993).

This chapter considers the fact that many children are still not learning and achieving at a high level, why NSF used the specific educational reform model that it did to enable all students to achieve at a high level, and what was learned from the previous implementation model if it were to be used again.

Our nation has decades of data demonstrating that most schoolchildren are not learning or achieving at high levels in our public schools. Internationally, we see that they do not perform well on tests of science and mathematics learning. When compared to much of the world, we rank in the middle, rather than in the top group of nations (Desilver 2017). In our national tests of reading and mathematics (NAEP), most of our children do not perform at an acceptable level. This poor performance is particularly true of the underserved minority children who now make up the majority of public school students. Ever since measurements were initiated, there has been a significant and persistent gap in achievement between minority children, many rural children and low socioeconomic status (SES) children (obviously, there is overlap in these categories), and middle and upper SES Caucasian children. Most significantly, the problem is growing worse, as many of the gains initiated in the 1990s are being lost.

This presents both an ethical and an economic problem. Given that there is much evidence that with the proper support and educational environment all groups of children can learn and achieve at a high level, it is simply morally unacceptable that a child's opportunity to succeed in life be dependent on an accident of birth. Also, there are well-established relationships between success in education and lifetime earnings, taxes paid, and social service cost saved (including incarceration costs). Our failed education system costs the United States billions of dollars each year. Since, as of this year, minority children make up the majority of the public school population and all credible demographic projections show that the population of underserved students will increase, this problem will only get worse if not addressed. Thus, it is clear that a massive reform of our education system is still required to create dramatic improvements in student learning and achievement until we reach the point that all students are well-served and reach the high levels of which they are capable.

The Urban Systemic Initiatives

The USIs and the systemic reform model were an attempt to support some of our largest urban school districts (the 25 urban school districts having the largest number of children in poverty were eligible for USI awards) in enabling all students to achieve at high levels. Systemic reform, as promulgated by NSF, depends on the understanding that a school district is a system in which the critical output is successful students. As a system, its components interact with each other through a series of positive and negative feedback loops, with delays in the flow of both material and information through the system. The interaction of these components, feedbacks, and delays generate the behavior of the system over time.

Envisioning an education system in this way has at least two important consequences in designing a reform intervention. First, it is necessary to identify the critical components of the system, the interactions and feedback loops between the components, the system boundaries and the optimal set points by which to measure the gap between actual function and optimal function. In this way, the levers that can drive reform may be identified. I suggest that NSF's seven drivers of reform (Curriculum, Instruction, Professional Development, Student Assessment, Policy Alignment, Funding Alignment, and Community Support) fulfill this first requirement; the requirement for the use of standards-based curriculum and instruction provides an interim set point against which to measure system change, while actual student achievement provides the ultimate set point used to measure system function. Second, the behavior of a system exhibits Liebig's "Law of the Minimum": growth (increased student success) is controlled not only by the total resources available but also by the scarcest resource (limiting factor). Only by increasing the amount of the limiting nutrient (the one scarcest in relation to "need") will further growth be possible. It then follows that increasing the limiting nutrient will increase growth to the point where some other nutrient becomes limiting.

The California class size reduction initiative of the mid-1990s was an example of Liebig's Law in education. The intent was to improve students' achievement by placing them in smaller classes. Besides mandating the smaller class size, the state even put up some funds to support it. If one thinks of small class size as a resource to improve student performance, then the intent of the initiative was to increase this resource and thus improve yield, or student performance. What actually happened was that by increasing the nutrient "small classes," another nutrient — namely, teachers — became limiting. That is, the number of teachers became the limiting resource and, in some cases, the number of classrooms also became limiting. This led to several negative consequences: large, inner-city school districts lost good teachers to smaller, upscale districts that also needed more teachers to meet the mandate of the initiative and could recruit them away from urban districts with higher salaries and better working conditions. Since the teacher pool could not be scaled up instantly, the urban districts were forced to use emergency certified teachers, actually weakening their workforce. Another result was that, in some cases, children wound up back in trailers for their classrooms.

I suggest that envisioning a system such as that proposed by NSF, and the role of Liebig's Law in its function, leads to the most efficient way in which to organize and operate an educational system to maximize student outcomes. It should be noted, though, that without sufficient resources (even if organized using the systemic reform model) the output of the system will be suboptimal.

After the implementation of the USI program in the mid-1990s, the school districts that won USI awards began to show significant positive gains. These gains were reflected in the widespread adoption of standards-based curriculum and instruction modalities, the large-scale training of teachers in the use of standards-based instruction, the teachers' increasing familiarity with the standards-based curriculum, the increased rigor of high school graduation requirements in mathematics and science, and increases in student achievement (several other papers in this group will discuss this).

Sadly, the NSF systemic reform approach has been replaced by a mixture of other reform models. Some I would describe as "market-based," in which a private company contracts with a school district to manage some or all of its schools; others function through awards to individual schools to be used to implement reforms; and yet another is the development of a significant charter school movement (groups of schools, either corporate or community run, or individual schools) in which a group receives a charter from the entity empowered to award it and receives public funding to manage a school or schools, but it operates independently of the school district. This charter school model has some characteristics of the first two models mentioned.

Current Education Reform and Comparisons with Systemic Initiatives

There are a number of reasons why the current mix of reform strategies does not work. First, the premise of the market-based approach seems to be that publicly run district schools have sufficient resources, but they are mismanaged, spending too many resources on administration (often including positions like school nurse);

the presence of teacher unions allows, or forces, the district to countenance a cadre of “lazy” teachers; and there is no competition. Parents have no real choice as to the public school their child attends. Thus, a “business-like” approach to school management will correct these deficiencies of management, labor deficiencies, and competition. I find it interesting that many people, who claim that schools are sufficiently funded and that throwing money at the problems of public schools, spend considerably more per pupil to send their own children to private schools, than what is spent per pupil in public schools. Further, given the nation’s recent history of having to bail out our largest manufacturing and financial institutions, I find this confidence in market-based solutions laughable. This approach also presumes that without unions, large numbers of teachers will be fired and replaced with a group of people who would become wonderful teachers if they were just paid enough. There is no evidence that such a willing group of great teachers exists and would be available. Further, this attitude, that more money is needed to get better teachers, is incredibly demeaning to our teacher workforce. While no one can argue that teachers, particularly new teachers, may feel underpaid (generally rightly so) and leave the system because of this, to posit that teachers working in the system purposely underachieve because of low salary is insulting. Last, school choice is a false solution given that many parents simply cannot move their children to other schools and that attractive schools have long waiting lists.

Second, the approach of making awards to individual schools is also seriously flawed. An individual school unsupported by a high achieving system is inherently unstable. For instance, the effect of changes in a relatively small number of personnel can dramatically alter the school’s environment. The literature is replete with examples of successful schools that have crashed after the loss of the principal or other key personnel. A single school cannot control or direct the district policies and resources on which it relies. Also, an individual school that has received an infusion of outside resources for a new program is dependent upon continuing infusion to maintain stability. It is often the case that once the outside funding is gone, the program it sponsored disappears. An examination of the kinetics of change for an individual school model argues against its utility. I explored this using a simple computer model to examine the kinetics of converting a system’s schools to effective schools. In one iteration of the model, I assumed that 80 percent of the district’s schools were not effective, and that in the first year of the reform 10 percent of the ineffective schools attempted reform. I further assumed that 75 percent of those schools were successful in raising student achievement and that 5 percent of the effective schools became ineffective each year due to personnel changes or other reasons. Then, in each following year, I kept the number of schools attempting reform the same and kept the other parameters the same. I found that after 10 years only about half of a district’s schools would be effective schools. In 10 years, approximately 55 percent of the district’s schools would be effective. If the loss of effective schools were 10 percent per year, rather than 5 percent, it would take significantly longer to reach even 55 percent (Burke 2007).

Third, we have the charter model. This appears to me to be a business-like model with the added feature that the school is funded by, but independent of, the school district. It has been argued that charter schools will promote school choice, become hotbeds of innovation, and force the district to improve in order to compete with charters, thus producing increased educational outcomes for charter students and leading to improved outcomes for all of the district's students. While there are some examples of success of these proposed outcomes, there is no significant evidence for any of these positive results being greater for charter schools than for regular district schools. Most studies of student achievement have demonstrated that most charter schools show no better outcomes for their students than public schools, and almost as many show negative impact on student test scores as those that show positive impact (Hull 2017). Increased school choice has been ephemeral because often the parents lack the ability to transport their children to charter schools or the time to donate to the school (often a requirement for enrollment). Charter schools argue that their enrollment procedures do not result in "creaming" of students; yet, the barriers to enrollment just cited and the fact that charter schools expel students at a much higher rate than regular district schools can or do argues that they do, in fact, cream.

I would posit that charter schools reflect all the problems inherent in the individual school reform model and raise some serious issues of lack of equity. Charter schools increase the already asymmetrical distribution of system resources, disadvantaging those minority and low SES students in non-charter schools. Inner-city districts generally suffer from a misdistribution of its most effective teachers. These teachers tend to migrate to schools with higher performing students, leaving the lower performing schools, almost always high minority/ low SES, struggling to get by with a teaching workforce that has a large proportion of younger teachers and a high teacher turnover rate, thus disadvantaging their students. How can anyone suggest that there is a shortage of effective teachers in a district faced with this misdistribution and, at the same time, suggest that clustering groups of the best teachers together in a single school will not do a disservice to the remaining students? In an ironic twist, it has also been observed that many charter schools suffer from a lack of effective teachers. In a district in which all of its schools are charters, one must argue that reshuffling a district's staff and resources in new configurations will somehow result in better outcomes. This is an essentially Darwinian model of school reform with the losers being many schoolchildren.

Last, the impact of student and teacher mobility on the system provides a further argument against reforming on a piecemeal basis. With student mobility rates of over 100 percent and teacher mobility rates in the 25–50 percent range, a checkerboard of schools differing in curriculum and instructional modalities cannot serve either the students or the teachers very well.

This mixture of reforms has not led to any real increase in student achievement and we are, as measured by test scores, either stagnating or regressing. Thus, I believe that our nation's educational system needs to implement a reform model based on NSF's systemic reform model. By building and growing the necessary standards-based infrastructure in all of a system's schools and aligning policy and resources to support this infrastructure, the system reform model is the model best designed to achieve long-term institutionalization of student success.

Lessons for Reformers

If future reform is to be based even in part on the NSF reform model, reformers should keep in mind the following lessons:

1. *Nothing is lasting unless it is represented as a line item in a budget.* Given this perspective, except for special circumstances, one should require that after three years any line item in the budget must be transferred to a line item in the systems budget unless the grantee makes a compelling case as to why it can disappear. We all have seen innumerable instances of good programs disappearing after the funding runs out, so pressure needs to be applied to institutionalize valuable components of the reform as it is being implemented.
2. *The leadership component of the reform needs to be strengthened.* While in the USIs we did insist that the project director be the district superintendent, the actual managing and delivery of the reform was in the hands of a deputy or other person. This is reasonable, but often that person did not have full authority to make commitments to institutional changes. Beyond the question of the deputy's authority in implementing and institutionalizing change, is the question of the job-life expectancy of a superintendent. Our experience was that friction often developed between a district superintendent and the school board, resulting in the superintendent being fired, being forced out, or resigning in less than five years. The new superintendent generally proposes different policies or practices than those of the ousted superintendent, leading to changes in direction that can harm the implementation of the reform. The school board needs to commit to the life of the plan that is originally proposed, modified only by changes that flow from the implementation.
3. *Reform needs to specifically focus on the production of a systemwide cadre of effective principals.* They are drivers of reform. Principals play a very important role in alignment of funding and policy. They coordinate all activities within a school for the common good. There was one USI site in which some of the schools had at least four different math reform awards at the same time. In addition, principal and parents interact frequently, particularly those parents who are opinion leaders. Reform-minded principals are able to overcome community opposition and bring the community on board in support of the program.
4. *Develop reasonable rewards for teachers engaged in the reform effort.* A school system is an ecosystem, rather than a mechanical system with only one goal. Generally, a stable ecosystem produces benefits for all of its populations.

Consider that in large-scale reform we ask teachers to assume a substantial share of the responsibility (even if unintentional) for the previous lack of success and ask them to change much of their behavior. While the reward of increased student achievement is a reward for most teachers, this reward takes a while to become evident. Systemic reform was often seen as just the next in a line of programs or requirements to which the teachers have been subjected, so serious attention needs to be paid to how to build their understanding of the process and how to increase their effectiveness as enablers of their students' achievement.

An aside: The power of the worker bees in a system became evident to me when I worked at the Center for Naval Analysis. While, theoretically the Chief of Naval Operations could have a sailor jailed for not following his wishes (program), it was clear that sailors often go their own way. Teachers, too, when faced with a new program (one of what is often a continuing stream) and, at the same time, are being judged by their student's test scores, can easily close their classroom door and proceed as they feel best.

5. *Timing,* a district needs to reform or rebuild the system starting from the bottom two grades and working up, prepare for implementation at the next grades, and provide extra support as necessary for the students in the upper grades who are not fully involved in the reform. It is tempting to suggest that there may be multiple starting points for systemic reform, but it is doubtful that there would be sufficient help to implement changes effectively across all the grades at the same time. While it might be useful to make some changes, such as beginning to implement a standards-based curriculum and testing across all grades in anticipation of coming reform, this would depend on the degree of cognitive dissonance engendered by the mismatch between the changes and what is already in place. The advantage of starting with the two lowest grades is that the kindergarten cohort will move into a reformed first grade and not lose ground. The first graders will move into a second grade that is moving in the right direction. This cohort would then move into the next two grades that have been prepared to implement the reform. Meanwhile, students, who will not get to benefit directly from all of the reforms, need extra support. There is no way for schools to help make up the 6,000 hours of enrichment that affluent children receive early in their educational careers that lower SES students do not without providing extra time, resources, and effort in their education. This can take the form of Saturday Academies, after-school programs, and so on.

6. *More sophisticated tools to understand the functioning and health of the system are required.* Since the function of systemwide reform is to increase the output of successful students, one could argue that this is the measure of the health of the system. However, it really doesn't function well as a guide toward building this health. Thus, we default to using a jigsaw-puzzle

approach to illustrate how system components interact. But this does not enable us to understand how the factors actually interact or to integrate the functioning of the separate components to determine how the program is functioning as a whole. Consider the difference in component versus system function. Depending on their relationship, if each of eight components of a system is functioning at 90 percent efficiency, the system itself may be functioning at less than 30 percent efficiency. To illustrate this, consider that when stacking blocks you only need to be off by a little in placing each block on top of the next for the pile soon to collapse. Thus questions arise, is the system effectiveness likely to increase? Is it likely to decrease? Will unforeseen outcomes arise? What will be the likely result if a policy or resource change is made? Systems models, whether equation-based (regression analysis) or agent-based, tell us a great deal about how the system is functioning and the likely impact of policy and resource-allocation decisions. They serve as a valuable guide in implementing systemic reform.

7. *The kinetics of process change, as it relates to reform, is critical to understanding and accountability* (Schneiderman 1988). While it would certainly be difficult practically and politically, the time period for implementing reform should be lengthened. In the early stages of real change, the situation may well become worse rather than better as the players learn to become effective in the new processes. In almost every field, one sees that process improvement, including the goal of raising achievement of all students, is not a linear process but follows some variant of exponential decay similar to the half-life of radioisotopes. In education, this means that if the achievement gap is cut in half in a given time period, the remaining gap will be cut in half, and so on. Thus, a realistic set of benchmarks for progress need be agreed on by all the important parties: funders, school boards, district administration, school administrators, instructional staff, and community.

In summary, school districts, like all entities, are subject to entropy and will decay without a constant input of energy. Reform, and the drive to enable all students to learn and achieve at high levels must be constant and ongoing.

References

Burke, D. D. (1992). System Dynamics-Based Computer Simulations and Evaluation: An Expert Anthology. In B. Williams and I. Imam, eds., Systems Concepts in Evaluation, 47–59. American Evaluation Association.

Darling-Hammond, L., Lieberman, A., & Miller, L. (1992). Restructuring in policy and practice. New York: NCREST Publications.

Desilver, D. (2017). U.S. Students' Academic Achievement Still Lags Behind That of Their Peers in Many Other Countries. Pew Research Center Fact-Tank. Retrieved from http://www.pewresearch.org/fact-tank/2017/02/15/us-students-international—math-science/.

Hull, J. (2017). How do charter schools compare to traditional public schools in student performance? Center for Public Education Data First. Retrieved from http://www.data-first.org/questions/how-do-charter-schools-compare-to-regular-public-schools-in-student-performance.

O'Day, J., & Smith, M. (1993). Systemic Reform and Educational Opportunity. In S. Fuhrman, (ed.), Designing Coherent Education Policy, 233–67. San Francisco: Jossey-Bass. Schneiderman, A. (1988, April). Setting Quality Goals, Quality Progress, 55–57.

Daniel Burke received the PhD in microbiology from Purdue University and served on the faculty in the Microbiology Department, University of Illinois; as chairman of the Biology Departments at both Mercer University and Seton Hall University; as deputy director for Education Research, CNA; and as director of the Doctoral Program in Computational Biology, New York University. His knowledge of the organization and function of K–12 education systems developed from his position as a program director in the USI Program and as senior staff associate for Systemic Reform, Directorate for Education and Human Resources, National Science Foundation, where he worked with all of NSF's systemic initiative programs. The insights he gained at NSF led to his use of system dynamics approaches and computer modeling to evaluate the impact of policy decisions and resource allocation on the function and effectiveness of educational systems.

Chapter 5

The Systemic Initiatives — Lessons Learned or Not Learned: A Personal Perspective

Margaret (Midge) Cozzens

The Last 25 Years of Reform

School systems grapple with all aspects that make up schooling: curricula and materials, professional development, assessments at all levels, the results of these assessments, parents who worry that their children are not challenged enough or worry that their children are challenged too much, a constantly changing group of administrators, and often students and teachers resistant to any minor or major changes. The systemic initiatives, particularly the urban and local systemic initiatives that focused on a single school system, confronted all of these aspects, often successfully and sometimes less than successfully. No matter how successful these initiatives were, however, sustaining change, even acknowledged positive change, has been hard. Administrators and teachers change, all in the context of new requirements mandated by governments at all levels, and including new standards in mathematics, language arts, and science, with mixed reception. Even though the students performed much better, not only in mathematics and science but also in language arts, in some cities — Detroit, Philadelphia, Miami-Dade, and Ames, for example — as a result of quality standards-based materials, competent well-trained teachers, and concentrated support of principals, superintendents, parents, and the community, we realized that students nationally across the board performed less well on international assessments and the National Assessment of Educational Progress (NAEP). The nation's solution — throw the baby out with the bathwater and start over again with new sets of "demanding" standards and assessments to try to prove that students in the United States are not really so bad.

This is not to say that the new standards are bad, compared to the old ones. What it says is that serious attention must be paid to **all of the component parts** needed for implementation of these standards, not just the standards themselves, or even the assessments related to the standards. We learned that much the last time new standards came to be. We have learned a lot from the era of local, urban, rural, and state systemic initiatives. Edmund Burke's famous quote, "Those who don't know history are destined to repeat it," applies to education as well as to world history.

On the positive side, we learned that standards-based instruction is not just a slogan, but that high quality curricular materials relative to the standards are essential and that teachers need to be prepared to use these materials. We learned that assessments are critical for the public view of education. Too often, results of state and national assessments drive the discussion about equity of educational opportunities independent of the lessons learned that assessments independent of curricular materials or teacher education and professional development won't work.

Evaluation documents give a number of reasons for the lack of success in some of the systemic initiatives. These reasons are attributed to people who tried to make the systemic initiatives work, as well as to those who did not want the systemic initiatives to work. These reasons often include not enough community support, changing administrators, lack of support for teachers, students who don't want to learn, disengaged parents, among many others. ***Little proof exists that there is even a full set of variables for success in education reform, much less that there is a set of minimum values for these variables.*** For example, Miami-Dade "pulled it off" despite the lack of community support and sufficient teacher support. What they did have that compensated for these things was a will to succeed no matter what and recognition that quality curricular materials could be used to drive nearly everything else, and in language arts in addition to mathematics and science. Also, and importantly, they were committed to sustaining reform past the early growing period.

Now let's consider the present. The Common Core standards in mathematics and language arts and the Next Generation Science Standards were developed with the business communities' interests and the college communities' interests front and center (Bidwell 2014). The belief was that if these two groups can get what they want from the change, they will get behind the change. This remains to be seen, but is there a commonality to college interests, much less business interests? We often hear from a company recruiter, "If only the graduates can think," yet thinking is often arguably not part of any state or national assessment. Colleges still use the mathematics ACCUPLACER exams, which require only a memory of often convoluted algebraic processes but no real problem-solving. Not much has changed in teacher preparation in colleges and universities either; future teachers are still taught content by faculty, who themselves were easily successful and expect the same from their students, and who lecture as the "sage on the stage" as they did 50 years ago.

The Overlap of My Own History with Reform

I left NSF as division director of Elementary and Secondary and Informal Education (ESIE) in 1998 after a conversation with Luther Williams, in which I told him that I thought the biggest single detriment to quality education in America and full functioning and sustainable systemic education reform was the colleges and universities across America. His response to me at the time was that I was probably right and that if I wanted to do something about it I had to go back into academia in an administrative role. Since I always did what Luther told me to do, I did just that. I went to the University of Colorado at Denver as vice chancellor for Academic and Student Affairs/Provost, a university with a large school of education in a large city. In 2000, “Before It Is Too Late,” a report by the National Commission on Mathematics and Science Teaching in the 21st Century, chaired by Astronaut John Glenn, called for (1) an ongoing system to improve the quality of mathematical and science teaching in grades K–12, (2) increases in the number and quality of these teachers, and (3) improved working environments for teachers. This report attracted a great deal of attention from government agencies, higher education, and others. The Mathematics and Science Partnerships (MSPs) programs were developed at the Department of Education and the National Science Foundation to provide funding to meet these goals. Funding for the MSP program continued until 2014, when it was discontinued. There is substantial evidence that at least some of the broadly defined MSPs have had a local effect on the directly impacted school systems, though these school systems are of modest size. Teachers whose content knowledge has increased as a result of summer programs and ongoing support will continue to be better teachers, but there is no hint of sustainability of any of the components of the successful MSPs impacting future teachers without continued funding.

Higher education national organizations launched their own efforts. I accepted a number of invitations to serve on national committees for teacher preparation. I was the only non-university president on the American Council of Education (ACE) Committee on Teacher Preparation. This committee produced an impressive report calling for university and college presidents to make teacher preparation a high priority on their campuses, to make it as important a priority as engineering and business and other areas. This compelling report gained next to no traction. I also served on a similar task force for the American Association of State Colleges and Universities (AASCU), which produced a similar report, winning a good deal of lip-service, yet also, like the ACE report, minimal traction nationally. One reason for this lack of effective response was the age-old “disrespect” between discipline departments and schools of education, each blaming the other for poor students and ultimately for poor teachers. Even at CU Denver, I was seen as a mathematician, not an educator, and one who did not know anything about education, though I voluntarily taught with the Rocky Mountain MSP.

On to business and industry! I then moved to the presidency of the Colorado Institute of Technology (CIT), not a university, started by the governor of Colorado to be the interface between business and industry and higher education in Colorado, believing that these groups did not communicate well with one another. With no state money in CIT, I raised money from businesses and companies in Colorado to incentivize needed changes in curricula to produce the future employees for Colorado companies. It worked — as an example, the many civil engineering firms in Colorado wanted or needed a specialization in transportation in civil engineering programs. CU Denver took the incentive and now has the first nationally ranked transportation program. Changing teacher education programs proved harder, despite calls from business and industry. Colorado was successful at creating very good Colorado State Assessments and these assessments did drive much positive change, change that continues today providing a quality education to urban, rural, and suburban students. The Rocky Mountain Math Science Partnership provided quality professional development, supplementing what was learned in colleges. Given that the governor was term-limited, and each new governor has his or her own agenda, CIT ended. Question: Can a teacher-variant of CIT be developed in states to facilitate all of the components of systemically improving K–12 education, in the light of the Common Core standards in mathematics and language arts and the Next Generation Science Standards, to secure funding from business and industry, rather than from the state or federal government?

Today

So what is happening today — we have new standards; two groups developing mathematics and language arts assessments, one of which is dominated by Pearson, a leading publishing company of K–12 books; and nearly no new curricular materials, although the first implementation of these standards occurred in spring 2015. Some of the quality materials developers of the 1990s into the 2000s are making revisions to their materials to emphasize the standards, but nothing appeared before spring 2016. Yet students are expected to take the new assessments in mathematics and language arts. Teachers don't understand what the process standards in mathematics even mean, much less how to prepare students for problem-solving relative to these process standards. Little professional development is occurring or is planned to occur relative to the new standards, and what is worse is many professional developers don't believe the new standards will ever come to pass, so they ignore them. Nothing is changing in mathematics, science, or language arts departments, or schools of education, so that teachers and students are able to learn using standards-based materials, whether they be new or old standards or new or old materials.

If this sounds pessimistic, it is not intended to be; it is meant to help people learn from the lessons of the past, especially those of the systemic reforms, before it is too late and we once again start all over. We need quality curricular materials and these materials must be prepared to help teachers learn what we all expect their students to learn. We cannot rely on colleges and universities to prepare the teachers for this new era and sets of standards, nor will there be sufficient professional development even planned to help these teachers. Yet, we do have examples of what works.

The Discrete Mathematics and Theoretical Computer Science (DIMACS) Center at Rutgers University has learned these hard lessons as we have developed instructional interdisciplinary modules in three different areas: biomathematics, computational thinking, and planning for a sustainable future. Modules in each of these areas "teach" content to the teachers as well as provide for students' learning. We have biology teachers in North Dakota teaching a full year course in BioMath to seniors, teachers who did not have an extensive mathematics background. We have mathematics teachers teaching BioMath in Raytown, Missouri, readily relying on the materials and biology teachers in their school to help them learn the biology. Watching high school students and teachers grapple with the computational thinking modules and the sustainability modules in summer programs indicates that teachers and students learning together has an enormous number of side benefits. One student said that she thought if she couldn't get the idea of the problem right away she must be stupid, but she saw the teacher struggle to understand it and gained an enormous amount of confidence.

The King Career Center in Anchorage, Alaska, offers career, vocational, and technical training for students primarily in grades 10–12. It is not a traditional high school since it focuses on career and vocational training (that which businesses profess to be interested in). Leesa Wingo, a lead teacher and field tester for the Computational Thinking project (VCTAL), teaches a veterinary science class at the King Career Center. After testing the 3-D Reconstruction Tomography module with her class, she offered the following comment: "One surprise for me — some of my quiet students showed serious ability in computational thinking. Their leadership skills appeared. The usually talkative students struggled, and the introvert thinkers thrived. The students enjoyed the spatial thinking involved in problem solving. On the end of the year evaluation for the class several students remarked about computational thinking activities being one of their favorite non-animal activities." Teachers' reactions to these modules indicate that they are able to learn the material and facilitate learning in their students. In Massachusetts, at the conclusion of the summer 2013 student workshop for Planning for a Sustainable Future (PS-Future), both of the lead teachers, one in mathematics and the other in environmental science, expressed their view that the modules met their basic goal of blending math, science, and sustainability topics in a way that was engaging to students. David Black (science teacher) said, "The modules … have the potential to serve many classrooms well. The integration of the math and science is clear, and the students responded positively to the applied nature of the curricula." In reference to the weather generator module, Jon Choate (math teacher) said, "The material served as a good intro to how global warming will affect the water cycle and the resulting changes in the weather. The material could be used in a variety of different settings: a science class, a math class, or a social science class. … The material was fun to teach and had some good hands-on exercises." ***When students and teachers learn, together everyone learns.***

Lessons Learned for Teacher Education

What does it take to design teacher education, both in-service and pre-service, so that teachers are learners learning as their students learn? Does crossing disciplines facilitate teachers as learners, and do we know enough about cognitive transfer of learning from one area to another? What does it take to make this change sustainable? These are hard questions, and I don't mean to imply otherwise, but I firmly believe that we have learned quite a bit from and since the era of systemic initiatives on how to better educate teachers, both before they officially become teachers and once they are teachers in on-going professional development. I will lay out four such lessons learned and address the impossible question.

First, we have learned that **locale makes a difference** and that locale is defined by the amount of money available to spend on education, socioeconomic background of the students and of the teacher, available community support, diversity within the schools, educational background of the parents, and politics. It means that teachers, new and old (most teach within 50 miles from where they got their degree) must understand the locale and that teacher education includes addressing locale or, said differently is *education in context*. The MSPs were designed to educate in context and this contributed to many of their successes, as did the successful urban, local, and rural systemic initiatives. The context of state initiatives was just too big in general.

Second, we have learned that **increasing the mathematical or scientific content knowledge of teachers independent of the availability of quality curricular materials doesn't work and that this content knowledge needs to be related to what the teachers teach**. This is a hard one since many academics believe that if the content is related to what the teacher teaches, it is necessarily dumbed down content. Not true! For example, if a middle school or high school mathematics teacher has always taught the expansion of polynomials like $(a+b)^2$ by having the students memorize the formula $(a+b)^2 = a^2 + 2ab + b^2$, and she learns that there is a geometric solution developed by the Greeks *Before the Common Era* that students can understand, both teacher and the students will remember how to develop the formula. Why not include this in a college algebra class, too? When we label courses as mathematics for teachers,

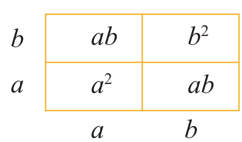

Figure 1. Expanding a binomial $(a + b)^2$

we assume the rest of the students shouldn't take them and that teachers somehow require more "elementary" mathematics. Since there are few new texts related to the new Common Core standards for mathematics or language arts or the Next Generation Science Standards, tying the content of courses that are a part of teacher education programs to these standards is more difficult but needs continuous attention. Those writing new textbooks should consider writing them so that teachers, as well as students, learn from using the textbooks. This also holds true for the revision of textbooks, such as *Mathematics for Elementary Teachers*, and others. These revisions seem to be on hold until everyone knows for sure what the new assessments are, but by then it may be too late.

The third lesson relates to learning as a progression from concepts to skills to problem solving to habits of mind and requires the **design of instructional materials and teacher education to help develop habits of mind**. Attention to concepts, skills, and problem-solving has been the focus to date of even quality materials and teacher support, especially done well in the MSPs and in the materials developed in the 1990s. Now attention is needed to the development of habits of mind. Related to this, most research on cognitive transfer has been on learning concepts, with little on more general problem-solving and habits of mind. The DIMACS interdisciplinary modules attempt explicitly to enhance students' ability to transfer ideas and strategies from one problem or contextual domain to another. For example, in the Heart Transplants and the NFL Draft module, despite the obvious significant differences between the NFL draft and heart-recipient contexts, on an abstract level the computational ideas and skills are identical. When a student exclaims, "It doesn't matter if it's hearts or quarterbacks, the numbers are the same," we know that ideas about making and assessing relative rankings have transcended particular contexts. But this cognitive transfer was a relatively easy step not requiring a larger analogical leap. Even so, not all students are able to take that short step easily. Working with students engaged in a variety of subject contexts as they attempt to think computationally has motivated the need for careful research on cognitive transfer (Belenky and Nokes-Malach 2012, 2013). More research and actual examples of effective transfer are necessary, along with a better understanding of how students engage in cognitive transfer.

The fourth lesson is that it is important to get teachers **to adopt a way of thinking that they are unaccustomed to** — not an easy task, nor easily sustainable, yet it is often, now and in the past, assumed to be automatic. "Show them the right way to think about something and they will think that way!" One reason for our success with interdisciplinary modules is that there are no previously held assumptions about how one should think about things outside a teacher's primary discipline and, working together, teachers from different disciplines better understand how others come to understand their discipline. Segregating those who expect to teach mathematics from those who expect to teach science, or language arts, or social studies, embeds commonly accepted notions, rather than opens students and teachers up to better understand how others think about their discipline and, indeed, themselves. An education course called *Teaching High School Math and Science*, or *Math and Liberal Arts*, would allow teachers to learn across disciplines with examples of problem-solving and habits of mind from many disciplines. Elementary school teachers have been expected to teach multiple disciplines, but they usually take courses specific to one discipline.

The last of the questions posed is the hardest and the one for which previous experiences in education reform provide no obvious answers: Is it even possible to create sustainable change, and does it make sense since education is certainly a dynamic enterprise? What we really want is to sustain positive change and eliminate negative change and to make systems responsive to externally imposed change, such as new standards, new administrators, new options (such as online learning), and new accountability measures. There are few lessons from the systemic initiatives, or the now discontinued MSPs on sustainability of reform. One example that still haunts me is from Ames, Iowa. In the late 1990s, the Ames school system, with a 4-3 vote authorization from the school board, agreed to order and then ordered new mathematics middle school textbooks, the *Connected Mathematics* series, which had already shown success in other schools. Extensive professional development had already taken place and administrators were committed to these new materials. In November of the first full year of implementation of these materials, the school board changed membership and voted, by a 4-3 vote, not to pay for the materials, even after the materials had been in use for three months. The school board was finally convinced after another three months that ethically they should pay for the new materials, so the students could at least continue using them and the instruction tied to them. Meanwhile, by year's end, the data collected on student achievement using the new materials showed substantial improvement for students at both the bottom end and the upper end of the spectrum compared to previous years when the students used traditional materials. The argument did not have to be made again. Creating sustainable change is hard and subject to a large number of different inputs, but we have to try and we have to persist.

We, too, as educators must be willing to change. Our students, like those in Ames, Iowa, deserve the best. My favorite quote from Maya Angelou's *Wouldn't Take Nothing for My Journey Now* is true today: "Each of us has the right and the responsibility to assess the roads which lie ahead, and those over which we have travelled, and if the future road looms ominous or unpromising, and the roads back uninviting, then we need to gather our resolve and, carrying only the necessary baggage, step off that road into another direction. If the new choice is also unpalatable, without embarrassment we must be ready to change" (Angelou 1994). I applied Angelou's words to education in a speech to the Margaret Warner School of Education graduates at the University of Rochester in 1996: "The 21st century will require each of us to repeatedly assess the roads ahead for education, and make iterative changes to provide all children and adults with the opportunities to learn. We must give our children the chance to assess and adapt to the roads they will encounter so that they make the right choices and contribute to our economy, society, and quality of life."

References

Angelou, M. Wouldn't Take Nothing for My Journey Now. New York: Bantam; reissue edition, 1994.

Belenky, D. M., & Nokes-Malach, T. J. (2013). Knowledge transfer and mastery-approach goals: Effects of structure and framing. Learning and Individual Differences, 25: 21–34.

Belenky, D. M., & Nokes-Malach, T. J. (2012). Motivation and transfer: The role of mastery- approach goals in preparation for future learning. Journal of the Learning Sciences, 21 (3), 399–421.

Bidwell, A. (2014). The history of the Common Core standards. US News and World Report. February 27, 2014.

The National Commission on Mathematics and Science Teaching in the 21st Century. (2000). "Before It Is Too Late".

Margaret (Midge) Cozzens is a Distinguished Research Professor and Associate Director for Education for DIMACS, the Center for Discrete Mathematics and Theoretical Computer Science at Rutgers University. She has served as Education, Outreach, and Communications Director for CCICADA, a DHS University Center of Excellence since its inception in 2009. She has managed a portfolio of over $10 million of NSF grants in education and evaluation over the past 10 years, including the development of high school modules in sustainability, computational thinking, and biomathematics, and for evaluation of undergraduate projects and the NSF VIGRE program. Her research is in mathematics and biomathematics, graph theory, game theory, and both science and mathematics education. Dr. Cozzens has over 100 papers and books in print, including a recent book, *The Mathematics of Encryption*, published by the American Math Society. Previously, Dr. Cozzens was president of the Colorado Institute of Technology and provost at the University of Colorado at Denver. She served for seven years in the 1990s as division director of, at the time, the largest division at NSF, the Elementary, Secondary, and Informal Education Division in the Education and Human Resources Directorate.

Chapter 6

Reflections on NSF's Comprehensive Regional Centers for Minorities (CRCM) Program: A Personal Essay

Costello Brown

Over twenty-five years ago (1991–92), Roosevelt Calbert, who was at that time the deputy division director in the Human Resources Division (HRD) of the Education and Human Resources (EHR) Directorate of the National Science Foundation (NSF), recruited me from California State University, Los Angeles, to NSF for an Intergovernmental Personnel Assignment (IPA). When I arrived, he told me that my first assignment would be to serve as the program director for the pre-college minority-focused effort titled the "Career Access Program," which included three programs, one of which was the Comprehensive Regional Centers for Minorities (CRCM). The CRCM had just recently been transferred to his division (HRD) from the Division of Undergraduate Education (DUE). It was probably transferred from DUE because the primary focus of this division was on undergraduates, while the primary focus of the CRCM was on underrepresented middle and high school students. At that time, the Career Access Program had a total budget of about $15 million. The other two programs in this pre-college group included Summer Science Camps (SSCs) and Partnerships for Minority Student Achievement (PMSA). Here is a brief description of the CRCM program:

> Supports developing systemic approaches at the precollege level to increase the presence of underrepresented minorities in science and engineering by establishing major Regional Centers. Eligibility: coalitions of colleges and universities, school districts, business and industry, professional organizations, state and local governments, and community groups. Dr. Costello Brown, Division of Human Resource Development (CRCM, 1991).

I had no experience as an NSF program director and, as a result, I depended very heavily on the constant advice and guidance of Roosevelt Calbert. The first thing that he told me was that I needed to site visit all the projects, and he gave me two consultants who he said could help me at each site. One was Lloyd Cooke, and the other was Eugene DeLoatch.

One of the major challenges with this assignment was that I had no idea what I was to do when I visited the centers: Calbert had just told me to "Go out there and see what's going on." I did just as I was told and, along with either Cooke or DeLoatch, visited all of the existing "centers." Most of the programmatic activity that we observed seemed to center on summer enrichment activities in mathematics and science for minority students, with a few projects having academic-year after-school tutorial programs in science and mathematics. Although some activities included parents, there was almost no focus on teachers and/or in-classroom activities. (More details are provided on this topic in a later section of this chapter.)

Subsequently, I updated the program announcement for CRCM with major guidance from Roosevelt Calbert; convened review panels; and added several new projects to the CRCM portfolio that focused more on teachers, classrooms, and school districts. In the current jargon, the two consultants (Cooke and DeLoatch), who went with me on the site visits, were really providing "technical assistance" to the CRCM project. In retrospect, I have no idea how Calbert knew that these two individuals would be so integral to the ultimate success of these projects and, at the same time, would provide the mentoring and guidance that I needed as a new and inexperienced program director. To my knowledge, neither of them had had substantial experience with K–12 schools and school districts. Eugene DeLoatch was the Dean of Engineering at Morgan State University for many years and Lloyd Cooke was a vice president of Union Carbide (a large chemical company) for 35 years. Cooke quickly became my mentor and taught me the most effective ways to provide assistance; successfully intervene; and bring resolution to often very complex issues with superintendents, principals, and other school district personnel. Cooke and DeLoatch were brought on board simply because Calbert told me to contact them because "[I] would need some help with the development of this program."

To summarize what was done: with the assistance of Cooke and DeLoatch, a level of sustainability was added to these projects by changing what had originally been enrichment projects (in which only the best and brightest students were selected as participants) to projects that now focused primarily on teachers and the classrooms. This was extremely difficult and some PIs strongly resisted since their projects had originally been funded as enrichment projects. However, we persisted and were successful to some extent, with the existing awards. The newly funded project did indeed focus on the classroom and teachers, as well as on students.

In later years, with the successor programs to CRCM, including the CPMSA (Comprehensive Partnerships for Math and Science Achievement) and USI (Urban Systemic Initiative), we were able to benefit immensely from technical assistance from Joe Harris and others in the McKenzie Group, led by Floretta McKenzie, who, among her other positions, had been superintendent of a large urban school district in Washington, D.C.

A Picture of CRCM

The picture would be considered relatively unique in many respects: it is unlikely that a similar picture of the PIs or co-PIs of any NSF program in the entire history of NSF would possibly have been more diverse — all but three of the PIs or co-PIs of the CRCM program were African American, Hispanic, or Native American. The program director and two consultants were all African Americans. No overt efforts were made to select this set of PIs, nor were there any outreach efforts made to encourage these individuals to apply to the CRCM program. Included would be deans, provosts, vice-presidents, STEM faculty, and a college president. It would be difficult to quantify the collective wisdom of this group of PIs, with respect to issues of equity, perseverance and educational role models. It was not by design and perhaps only a lucky coincidence, but I could not have asked for a better and more rewarding first assignment than as a program director at NSF.

I have taken the liberty of listing the names and affiliations of most of the PIs or co-PIs in the CRCM program because of the tremendous contributions that they made, not only to the success of the CRCM program but also to the improvement of the teaching and learning of mathematics and science for the K–12 students and teachers in the school districts that they served: Jewell Cobb, former president of California State University, Fullerton, was the PI on the Louisiana CRCM project; Eric Hamilton and Dorothy Strong were the PI and co-PI of the Chicago CRCM project; Melvin Webb was the PI of the Atlanta, Georgia, project; David Blackwell, engineering faculty at Northeastern University, was PI of a project that worked with several schools in Boston; Patrick Weaselhead from Montana State University was the co-PI of a CRCM project that focused on American Indian students; Alfredo de los Santos was a vice president at Maricopa Community College and the PI of the Maricopa CRCM, which focused on Hispanic students. Ernesto and Nora Ramirez also worked with the Maricopa CRCM; Harvest Collier, the PI of the St. Louis CRCM, is engineering professor and, more recently, retired vice provost for undergraduate studies at Missouri University of Science and Technology; and, finally, Vicente Llamas, physics professor, was the PI of the New Mexico Highlands University CRCM, along with Betsy Yost, the co-PI, whose program targeted Hispanic and American Indian students in rural New Mexico schools.

In my first year at NSF (1991–92) my immediate boss (Calbert) as well as the assistant director of the EHR (Luther Williams) and the director of NSF (Walter Massey) were all African Americans. In other words, except for Joe Danek, HRD division director, the reporting lines all the way from me, as a program director to the director of NSF were all African Americans. This specific time in NSF history, with these specific individuals, provided a context and a fertile ground for the development of CRCM, as well as all of the systemic initiatives and other equity-focused programs in the HRD. All of us, from the NSF director, to the program director (Massey, Williams, Calbert, Brown), were born in the south, attended the "separate but equal" segregated K–12 schools, and graduated from historically black colleges and universities (HBCUs): Massey, Morehouse College; Williams, Miles College; Calbert, Jackson State University; and Brown, Hampton University. Three other program directors in HRD at the same time (Roosevelt Johnson, William McHenry, and J. Arthur Hicks) would all fit this same profile. As a side note, Williams graduated from Hale County *Training* School in Alabama, and I graduated from Caswell County *Training* School in North Carolina. One could only imagine the time and energy that went into denoting the black high schools as "training schools" throughout the South during that era. We all had a personal understanding of poverty, discrimination, and racism, and all the baggage that it brings to an educational environment. We had all succeeded educationally and received PhDs in STEM disciplines, in spite of all the obstacles presented to minorities in the South throughout the 1950s and 1960s.

As mentioned earlier, in 1992, NSF revised the mission of its Career Access program and changed its emphasis from undergraduate level programs to pre-college/K–12. This resulted in the move of the Career Access Program from DUE to HRD. I would submit that without this critical decision (presumably made by Massey and Williams) to make K–12 education a higher priority for NSF, it is unlikely that the events that followed would have occurred, at least not in this time frame. However, in all likelihood, the wheels for the new focus on K–12/pre-college and the severe underrepresentation of minorities had already been set in motion by Eric Bloch, the eighth director of NSF, who was the director when Williams came to NSF in 1989 as the director's senior science advisor. A year later, Williams was appointed assistant director (AD) of EHR and Massey succeeded Bloch, as the ninth NSF director, beginning his tenure in March 1991 (Massey 1992).

Modifying the Traditional Role of NSF Program Directors

I didn't realize it at the time, but as NSF program directors in the HRD, we behaved in a somewhat intrusive manner with PIs. I took my cues from Williams and Calbert and initially thought this was normal, and the way all program directors at NSF operated. In those days, there were often multiple site visits to the same project at the first sign of any problem. Some issues with schools or school districts were nonnegotiable. For example, after one site visit to a school district, I came back and reported that we had numerous instances of "a school within schools," which simply meant that the schools were integrated, but almost all the classes within the school had either all black students or all white students. I had no difficulties and needed very little explanations and/or clarification before getting sign-offs on withholding or, in some cases, denying funding for proposals that had reviewed well (the data was indeed integrated) but only a site visit revealed major equity concerns and the fact that the students themselves were not integrated. I was really surprised to see that some districts refused NSF support when they realized that the funding would be conditional and the project had to include all children.

I realize in retrospect that EHR and, more specifically, the CRCM, really did not fit well within the usual NSF paradigm. As a PI on a regular NSF research grant, I would be evaluated primarily by my publications or patents. Over the course of the three-year award, it would be highly unlikely for an NSF program director to site visit me and then start telling me to change my approach or to interfere in any way with my stated plans in my original proposal. If I made satisfactory progress in achieving my stated goals and objectives, I could apply for and possibly receive a three-year renewal and continue business as usual. The culture of NSF strongly reflects a perspective of the academy; we are collegial, we do scholarly work, we write our proposals, we get our grants and carry out our scholarly work, and no one (especially from NSF) interferes. The only sense of urgency is that someone else might scoop us and publish our findings ahead of us.

The CRCM began to implement a different culture of accountability, supported by formal technical assistance, and elements of this culture were greatly expanded and magnified by the Urban Systemic Initiative (USI) program. The era of clearly specified objectives and deliverables with time lines monitored by somewhat *intrusive* NSF program directors had begun.

Unfortunately, there were no CRCM proposals being submitted from the most severely underserved school districts, the ones that needed it the most, and even some of the cities that were awarded CRCM grants had dismal statistics/academic achievement data that on the surface are hard to believe. For example, I site visited a high school in St. Louis, where I was told that in the previous five years, there had been only one graduate of this high school that had gone on to college. At that time, the perceived better students were being bussed to suburban schools.

Overview of the CRCM Program

No cookie cutters were used to develop the various CRCM projects and they were all quite different in their design. However, the overarching theme and focus was K–12 (mostly middle and high school) minority students and mostly from large urban school districts, including selected schools in urban school districts: Boston, Chicago, Los Angeles, Atlanta, St. Louis, Philadelphia, and Phoenix (Maricopa County, Arizona). In addition, two rural areas were included: American Indian students from several reservations in Montana and Hispanic/American Indian students in several rural New Mexico school districts.

Some of the CRCM projects, such as the CRCMs in New Mexico and Atlanta, implemented relatively rigorous summer enrichment programs in science and mathematics on college campuses. In most cases, the middle and high school teachers were not part of these activities. Probably half of the projects initially focused on the students (almost all participating students were underrepresented minorities); however, some projects, such as the ones at Maricopa and New Mexico Highlands, focused primarily on support for teachers and their STEM-related classroom activities with the teachers actively involved.

The CRCM program slowly evolved, with some gentle encouragement and pushing (by an intrusive PD), to a focus on classrooms and teachers where it was clear that there would be greater impact and the impact could be sustained much longer than by just solely focusing on students. In the early days of the CRCM implementation, site visitors seldom met with superintendents or principals. As the programs evolved from CRCM to USI (Urban Systemic Initiative), site visitors met mostly with superintendents or deputies and building leaders, since now the focus was broadened to bring about district-wide reform in science and mathematics. Several program directors for the USI later told me that the cities that seem to have had the most success were the ones that had originally participated in the CRCM.

Prior to his departure from NSF in 1999, Williams had created the new Educational System Reform (ESR) Division in the EHR and appointed me as the acting division director. After a search, I was appointed as the division director of the ESR. Ultimately, CRCM and PMSA (both part of the original Career Access Program) became the CPMSA (Comprehensive Partnerships for Minority Student Achievement), which was moved over to ESR. The CRCM/CPMSA served as a precursor for a new Urban Systemic Program (USP). An attempt was made to include all the lessons that had been learned from CRCM/CPMSA and the USIs in the program announcement for the new USP program. Celeste Pea, who had worked extensively with the USI, was a major contributor to the initial USP program announcement. In addition, the Rural Systemic Initiative (RSI) was also born and housed in the ESR.

One can get really confused by all the acronyms in the above paragraph, but in essence what it says is that I was now in charge of a division that had both the urban and rural systemic K–12 portfolio with a budget of approximately $200 million. The program directors included Celeste Pea, Julio Lopez-Ferrao, and Kathleen Bergin, who worked primarily with the CPMSA and USP projects, as well as Gerry Gipp and Jody Chase, who co-managed the RSI projects. I list the names of these program directors because I would submit to you that to the extent that these programs were successful was due in no small part to these individuals. For sure, the rhetoric has been somewhat muted, the word “systemic” was removed from NSF jargon for a while; however, accountability and proactive interactions with PIs, remained as fundamental guiding principles.

We had a slogan in the ESR that we should always remember: “Sam and Samantha.” In other words, never forget that ultimately it is all about the students, not egos, not program directors, not PIs. Numerous evaluation studies have measured the effectiveness and impact of the aforementioned K–12 programs in the ESR and that could easily be the subject of another essay. It would be difficult to quantify the contributions the USI made to the USP or to the RSI. Participating Native students were now engaged with curricula revised to include Native culture and “Native Ways of Learning”; students and teachers in remote Appalachian regions of Kentucky and Ohio were now active participants in hands-on science and mathematics. Wimberly Royster, as PI of an RSI project and subsequently as PI of a Mathematics and Science Partnership (MSP), project was an excellent example of PIs who benefited from their experiences with the USI, RSI, and/or closely related programs. Royster's projects had a significant positive impact on math and science achievements for teachers and students in the Appalachian region.

Significant Contributions

Finally, in my opinion, one of the single most significant contributions that Luther Williams and Roosevelt Calbert made to the CRCM and other K–12 initiatives at NSF was that a system that promoted a strong sense of urgency was created on the part of grantees (PIs, superintendents, other educational leaders) that I had not observed before the SI era. This seemed to have been done quite adroitly through a relatively unique appeal to the basic human emotions of fear and pride, as well as to public accountability (those dreaded "midpoint reviews" and reverse site visits). It was during this period that seldom did I hear the name "Luther Williams" mentioned without a few adjectives in front. However, a very important message went out to the education community at large that screamed loudly: (1) This is not business as usual; (2) SHOW ME THE DATA; and (3) Please don't tell me what you plan to do, when you should be telling me, rather, what you HAVE done. At the same time, more subtle messages were also being sent: (1) This work is of national importance (so why shouldn't the AD be in attendance at the midpoint review and lead the discussion?); (2) Too many generations of inner city kids have already been *under- or mis-educated* in science and mathematics. How many more generations have to be sacrificed to the gods of politics and research findings that are never implemented? Something, indeed, is wrong with this picture.

The primary reason for my interest in reflecting on the CRCM, and the USI era at NSF and its overall contributions to the education of students in mathematics and science in urban and rural school districts is that, hopefully, these efforts may shed some light on the creation of more promising successful initiatives focused on math and science for underserved students. As Shirley Malcolm often says, "These problems don't stay solved." The intrusive/interactive, but supported, cooperative agreement model was indeed somewhat alien to the culture of NSF; however, it was incorporated, to some extent, into a subsequent K–12 MSP program. (I headed the first two committees that developed the initial MSP program announcement). We have learned from experience that such a multifaceted approach is needed or required to even begin to provide a sustainable quality K–12 education in an environment of poverty. We have also learned that a culture of equity in an increasingly multicultural environment requires constant attention and work.

In my humble opinion, harnessing the collective wisdom and experiences of the folks engaged with the CRCM program went a long way toward addressing the educational achievement gaps and disparities that seemingly have persisted forever. I will end this chapter with a quotation from *The Prince* that Luther Williams used in concluding his plenary remarks at an EHR Diversity Conference in 1992 (Machiavelli 1515):

> As the custodians of the future, we must be willing to make the effort, remembering all the time: "There is nothing more difficult to carry out nor more doubtful of success, nor more dangerous to handle, than to initiate a new order of things. For the reformer has enemies in all those who profit by the old order, and only lukewarm defenders in all those who would profit by the new order."

We must all remember that, ultimately, "we make the effort" for Sam and Samantha and for all the other underserved students in mathematics and science that they represent in our K–12 schools.

References

Comprehensive Regional Centers for Minorities (CRCM)/Career Access. Brochure NSF 91-129. February 14, 1992.

Machiavelli, N. The Prince. Written c. 1505, published 1515. Translated by W. K. Marriott.

Massey, W. (1992). A success story amid decades of disappointment. In E. Culotta and A. Gibbons (eds.). Minorities in science: The pipeline problem. Science. 258: 1177–80. Proceedings of the National Conference on Diversity in the Scientific and Technological Workforce, September 25–26, 1992. Washington, DC.

Costello L. Brown is professor emeritus of chemistry at California State University, Los Angeles. He is a former associate dean and acting dean of Graduate Studies and Research at the same institution and has worked as a senior associate and consultant with the Quality Education Minority Network (QEM) for over 15 years. He worked at the National Science Foundation as division director, Education System Reform, which made systemic awards to both urban and high poverty rural K–12 school districts, as well as to tribal colleges under the Rural Systemic Initiative (RSI) program. At NSF he also was a program director for the Alliances for Graduate Education and the Professoriate (*AGEP*) program and helped to develop the Mathematics and Science Partnerships (MSP) program. He has served as a QEM consultant/mentor for two Leadership Development Programs focusing on faculty and staff at tribal colleges with NSF Tribal College and University Program (TCUP) awards. He received the PhD in organic chemistry from Iowa State University of Science and Technology and a bachelor's degree in chemistry from Hampton University. Throughout his career, he has worked to increase the participation of underrepresented groups in science, engineering, and mathematics. He has served on a wide array of review panels and study sections, both at NSF and the National Institutes of Health.

Chapter 7

Technological Innovation and Urban Systemic Reform: Designing for Change

Jere Confrey

> Systemic reform is not merely an issue of producing more of the innovation. It requires attention to organizational issues of the reform and the research and development enterprise to learn about creating sustainable innovations. —Blumenfeld et al. 2000

Since the implementation of the systemic initiatives (SIs), we watch with dismay in this country at the state of disregard for, or — worse — increasing hostility toward, the pressing needs of children, especially in urban settings. The situation breeds hopelessness, blaming of victims, and an intolerable acceptance of failure. Interventions *are* required and would greatly benefit from being informed by what was learned from the systemic initiatives. During systemic reform, we saw the emergence of increased capacity to drive change toward implementing high standards for all and, afterward, these resources in many places have diminished and have been lost.

The question I selected for this reflection is "What did we learn from the USI experience that provides tools and a roadmap for the role of technological innovation to create positive change when practiced equitably and sustainably?" I had the honor of being on the review panels for 26 states' SSIs and for many of the USIs. I also worked as a member of technical assistance teams, visiting, meeting, and working with many of the USI cities. My ongoing interest in systemic reform is currently informed by recent work on envisioning the next generation of technology for the digital delivery of curriculum with scaffolding to support a focus on big ideas, learning trajectories, and diagnostic assessments, and taking such work to scale. Juxtaposing my recent work in the technology-based subsector of the education industry, my current work designing new forms of digital-based assessment and learning maps, and my work in technical assistance to the systemic initiative program (urban, rural, *and* state) have been the crucible for the following reflections.

To begin (with the benefit of hindsight!), I believe that three components would have greatly improved the outcomes and sustainability of the SI programs had they been required prior to an SI's funding award and activation. These amount to a set of preadaptive conditions for enhancing the probability of major strides and success of SIs. Instead, these conditions were often themselves the de facto goals (or necessary accomplishments for the goals) of the SIs:

1. Sufficient technology: with widespread access, reliable and fast connectivity, and consistent maintenance support;
2. Compelling materials and tools; and
3. Sustained professional development.

Post-systemic reform, we recognize that compelling materials and tools are critical. During the systemic reform era, examining and recommending stronger curricula was common, but requiring its use was outside the enforcement of the agency. Many sites did take advantage of some of the NSF-supported curricula, but they were seldom able to offer sufficient professional development to make its implementation successful. The call for this is not a criticism but a suggestion for moving forward, informed by the site visits. I will also draw on some related efforts that did entail both systemic reform and novel implementations of curriculum.

Critical to this call for these components is the recognition that it is the *interplay* of the components, or the *engineering* of their interactions, that is important. For many years, I have referred to a metaphor provided in an essay called "Oppression" by Marilyn Frye, a feminist scholar. She used the metaphor of a birdcage in discussing oppression:

> If you look very closely at just one wire in the cage, you cannot see the other wires. If your conception of what is before you is determined by this myopic focus, you could look at that one wire, up and down the length of it, and be unable to see why a bird would not just fly around the wire any time it wanted to go somewhere. ... It is only when you step back, stop looking at the wires one by one, microscopically, and take a macroscopic view of the whole cage, that you can see why the bird does not go anywhere; and then you will see it in a moment. (Frye 1983, p. 3)

It is only with the recognition of the configuration of *all* the wires that the concept of a cage becomes a valuable metaphor, complete with its oppressive connotation. I introduce this metaphor into the discussion at this point to remind us that the components listed above interact with each other, and their absence or insufficient functioning, for our children, serves to maintain multiple wires of a cage of constrained learning and impaired opportunity. The systemic initiatives were a remarkable and an unprecedented attempt to remove multiple cage wires at a time and find ways to create the more interactive systems needed.

For effective interventions at scale in a school or district, the school's (or district's) current conditions need to be examined through a systemic lens. For example, **access to technology** in schools is clearly a required component, but establishing fair access in schools does not tell the entire story. The degree of general prior knowledge and experience students and teachers bring to tasks, the inclination and habit to look for related resources, and the appearance of competence in school all play a part in systemic reform (Wenglinsky 2000). Differential access at home also must be considered: home access is a critical issue that may determine a student's success on a project or throughout a course. Experience in systemic reform provides an impetus — and, I believe, a kind of wisdom that informs us — to look harder at the meaning and significance of how the components function at the system level

But before we discuss design elements of the interplay, a few comments on the current state of these components in the post-USI world seem useful. I first consider aspects of the three conditions listed above, and follow with comments about the relationship between innovation and systemic reform, and the notion of engineering for change.

Technology

Data from the National Center for Education Statistics (Gray et al. 2010, from NCES is the most recent national data) showed that in 2008 there were 3.1 students per computer in schools on average. This ratio describes the number of computers per building but does not report how many computers are actually available to students on a regular basis, nor does it consider how the computers interface with classroom instructional practice. In many schools, one observes that computers still reside, for the most part, in labs, and while slightly better access is afforded with carts, computer resources remain marginal in the daily practice of schools. Furthermore, in the world of new technologies, issues of connectivity, networking, and high-speed reliable access persist, especially as more and more curricula live in the cloud and not on the individual computer. With the advent of tablet computing and low-cost notebook computers, such as the Google Chrome notebook, the demand for technology in schools, and multiple local and statewide initiatives to expand schools' broadband access, the situation has no doubt improved considerably in the past nine years. But the use of the technology to support learning and reasoning remains a key concern.

Drilling down, one sees the expected digital gap emerge in urban settings. Too many urban settings use computers for drill and practice, and they leave students alone with apps that reinforce a view of cognition as individual acquisition of knowledge; computers are often used to manage students rather than to teach them. At the same time, we realize that the world for which students should be prepared demands from them "literacy in information and communications technology (ICT) — which relies on skills such as thinking and problem solving, communicating effectively, self-direction and productivity — requires fully integrating technology with classroom learning" (NEA 2008). This observation still holds true today. Using technology to achieve real learning goals is strikingly absent in most urban settings. Exceptions such as Union City, New Jersey (Honey et al. 2000) remain strikingly atypical.

The “easy” solution is to apply a version of deficit thinking and call it innovation — treating children as damaged goods, spoon-feeding them materials below grade level, and tricking them through games into practicing and, presumably, learning. While clever forms of strengthening students’ “readiness” can be a positive force, a steady diet limited to remediation is no service at all. If technology is used to reinforce *low* expectations about children’s ability to learn, then it *contributes* to the problem of accelerating and strengthening the education of underserved urban (and rural, of course) children, rather than promoting a durable solution to this dilemma.

Curricular Materials

If the Common Core State Standards (CCSS) — or close analogs adopted by states in the name of “rebranding” the CCSS — hold in the current unreasonably polarized national political and policy environment, then it is possible that all students will be expected to achieve ambitious instructional goals. The extension of these goals to all students was clearly afforded by attention to systemic reform (O’Day and Smith 1993) and via Title I fund authorizations (U.S. Department of Education 2016). The fundamental principle of high expectations for all students’ learning is in part a tribute to both the audacity and the determination of the systemic initiatives.

Because NSF supported the creation of reform materials in the 1990s, the USIs were offered, and often even counseled, to implement curricula that were aligned with high standards. Regrettably, NSF suspended its curricular agenda during the post-systemic reform period, at great cost to U.S. students. Abandoning, to commercial developers, the development of materials designed to meet Common Core State Standards has resulted too frequently in the production of rushed and mostly cosmetic changes to existing textbooks. This situation’s occurrence was partly due to the lag time between curriculum development and the release of standards; it may be remedied over time as some publishers move toward “real standards-based” materials. The Common Core sponsoring organizations (CCSSOs and NGA) knew to avoid curricular discussions: they sought to have CCSS — the actual learning expectations — adopted by themselves. They tried to avoid political charges that the standards were a national curriculum, or a Trojan horse for a national curriculum, which, of course, they were not “These Standards do not dictate curriculum or teaching methods” (Common Core State Standards Initiative 2010, p. 5). However, a nationwide, scattershot hodge-podge of materials, including the use of open-source free web materials was created by states and districts, and they weakened the implementation of Common Core. This left parents, who were often themselves uninformed, at sea because they failed to distinguish between the goal statements of the standards on the one hand, and curriculum, which was much slower to emerge, on the other. This was often massively exacerbated by deliberate misinformation by some groups spearheading various forms of backlash against the Common Core and the associated assessment initiatives.

The way contemporary publishing is carried typically decreases the quality of the materials. Short turnaround times, immediate market orientation, minimal testing of materials, lack of sufficient attention to outcome measures, and aggressive publishing-house marketing and production are papered over by the appearance of distinguished author teams, while still leading to inferior products. However, some reform materials developed during the systemic reform period have persisted and even succeeded, and they are being revised toward CCSS goals (Everyday Math, Connected Math, Math Investigations, and Core Plus), but it does remain a question whether revisions are sufficient to the challenge of the new standards. Whether efforts to improve or update curriculum by adding technological resources really drives a sufficiently innovative approach to curriculum remains to be seen.

Professional Development

From the local, urban, and state systemic reform efforts, we learned firsthand how critical it is to have well-qualified, middle-range capacity professionals in the content areas in schools and districts. Where a number of devoted, well-qualified teacher educators and coaches emerged and were supported by districts in, for instance, Phoenix, El Paso, Detroit, Montana, and Louisiana, to name a few, we witnessed progress in urban or state systemic reform. A layer of dedicated curriculum specialists was critical to stability and to sustaining the reform going forward (Weiss et al. 2003). This group of curriculum specialists, with long-term commitments to teachers' professional growth, concentrated their work on the instructional core (City et al. 2009; Elmore 2000). Superintendents mattered, of course, for considered leadership and wise decisions, but regrettably the continuity of superintendent-level leadership was undercut in many localities by frequent turnover.

However, the Great Recession burned its way through the U.S. economy in 2007–9. Local, regional, and state-level Departments of Education or Public Instruction shed curricular specialists at an alarming rate. A state might be left with one or two experts per grade band, struggling to meet the needs of change required by new standards. (I saw this in North Carolina, for example.) Partly as a result, the professional development around CCSS devolved too often into a superficial introduction to the structure of the standards, rather than into an in-depth analysis of the content and practices, especially as was necessary for informed and competent implementation in classrooms. Rather than tackling the deep content implications of, for instance, intensifying statistics in middle grades, or the requirement of at least one-third of a year of algebra for all students by the end of eighth grade, professional development was generally limited more typically to broad and generalized introductions of strands, recitations on practices, and, by emphasizing coverage of standards one by one, a pattern of "teaching to the standards."

Second, the move under the No Child Left Behind Act to link teacher accountability to their students' performance on standardized tests came to be perceived as a premature and punitive quality that undermined the professionalism of teachers' work. While the Gates Foundation (2010) produced sound and thoughtful advice on the implementation of teacher evaluation systems, this advice has been largely ignored. Instead of becoming a form of professional communication and plan for improvement as recommended, Common Core implementation efforts frequently comprised a caricature of a systemic approach, lacking high quality, adequate opportunities for improvement, multiple student outcome measures, or the use of sensitive analysis of data.

Understanding the Relationship Between Innovation and Systemic Reform

Systemic reform at NSF in the 1990s focused on a common set of features: (1) standards, curriculum and materials, professional development, and assessment; (2) convergence of financial resources; (3) convergence of policies; (4) partnership among in- and out-of-school programs; and (5) data on outcomes. Successful applicants were obliged to show how they brought together these features and, in doing so, how they addressed the crosscutting issues such as equity and technology. These factors proved to be essential to creating the conditions to support change and innovation, as witnessed in a variety of the SI programs.

For example, in Detroit, a confluence of NSF-related work from systemic reform and advanced technologies permitted the development of a unique experiment. Faculty from the LeTUS Center (Center for Learning Technologies in Urban Schools), many from Northwestern University and the University of Michigan, joined with leaders of the Detroit Public Schools to implement a technology-rich project-based middle school curriculum. Their innovations comprised creative materials, professional development, and technology use. Reflecting on the experience, they identified three critical systemic factors: (1) policy and management, (2) professional development and human capital, and (3) technological capability (Blumenfeld et al. 2000). There are lessons to be seriously considered when determining how to drive innovation forward to meet the current needs of similar urban schools.

1. *Policy and management.* A recognition emerged that curricular change was often impeded if the policies of the district were not aligned from bottom to top with the reform. Buy-in by administrators was not synonymous with the buy-in by classroom teachers. Competing reforms often overburdened teachers.

2. *Professional development and school culture.* The innovation team realized they had underestimated the demands placed on teachers for planning and overestimated teachers' willingness to do the planning, especially given their many pressures and priorities. Strict policies, related to the collective

bargaining agreements of the school system with the teachers' union, often made it very difficult to plan, prepare, collaborate, and build a professional community of practice during after-school time. There was a general lack of teacher preparation for fostering inquiry and related classroom management issues. "We anticipated some tailoring to circumstances, but were not prepared for the difficulty of helping teachers adapt at the same time that they were learning how to enact new practices" (Blumenfeld et al. 2000). Furthermore, they acknowledged a much deeper need for change in professionally serving teacher's needs.

Ambitious instructional reform requires a school culture that supports professionalism and provides opportunities for sharing, risk taking, and reflection among teachers about pedagogy and student learning. It is also more likely to take root when there is open communication and cooperation among administrators and teachers about what works and what is needed so that policy and practices that support such pedagogy are not only established but reinforce each other as well. An organizational culture that emphasizes students, rather than bureaucratic procedure and dictates, is also critical. In reflecting on what would be necessary to truly support systemic reform wrapped around curricular and technological innovation, they identified the need for rich video databases and online communities. In the nearly two decades since the work of the LeTUS initiative (and many others), our research-based understanding of the necessary conditions for teacher change and systematically improving the instructional core has become much richer.

3. *Technology capability.* The studies of technological innovation within systemic reform initiatives spoke to basic needs for technological infrastructure. They recognized the need for teacher competence at a time when only 23 percent of teachers used computers at home. In addition, innovation was hampered by a lack of reliable access to the Internet and a sufficient number of machines. Though systemic reform in the 1990s did not rely heavily on technological innovation in urban settings, the Detroit experiment (Hug et al. 2005), linked with a number of the lessons from an experiment in Union City, New Jersey (Chang et al. 1998; Honey et al. 2000), provided evidence that in the future, these kinds of technological innovations should be coupled with a systemic perspective when addressing the needs of urban youth.

Engineering for Change

What lessons should be taken away from the hard-won insights that the systemic initiatives and the learning sciences made possible? What changes are needed to address these insights and engineer more successful progress going forward?

One insight is that simply importing innovation wholesale from the lab or other settings is highly unlikely to be transformative without careful attention to systemic factors. Cohen and Ball (2000) speak of a gap between specification and development, on the one hand, and implementation, support, and maintenance, on the other. They suggest that to expect that one can begin by defining what is desirable, then predicting how to get there is likely to result in expectations that fail. For example, when teachers implemented project-based science, they tended to emphasize familiar units, while avoiding or shortening new units, thus undercutting the meaning and potential impact of the innovation. Furthermore, they tended to assign reflection questions only as homework, to have difficulty bringing the innovative curriculum to closure day by day, and ending sessions at the completion of activities instead of what we now recognize as necessary steps after the activities, namely, reflection and closure of the *learning experience*. The authors found that adapting new practices required much more time and effort than teachers could devote to the innovation (in the context of the various constraints on teachers when the innovations were introduced). Without careful consideration and partnership around issues of implementation, support, and maintenance, reform and innovation are, simply, unlikely to succeed:

> This dynamic [innovators pushing innovation and schools pushing to conform to norms] is congruent with the caution Stokes, Sato, McLaughlin and Talbert (1997) made that the challenge of getting a reform to scale is more than adding numbers of classrooms. It is the challenge of translating an externally conceived and supported program into one that will be internalized so that it is conceptualized, governed, and practiced in the schools. In this way, ownership shifts and underlying norms and practices change to sustain the reform principles. (Blumenfeld et al. 2000, p. 159)

Even this advice may prove insufficient. Tough decisions must be made early, regarding which critical but lean factors to include and the intentional creation of deliberate (engineered) ramps to implementation. Incremental implementation, together with a willingness for active and frank reappraisal and revision, may be essential to learning painful but important lessons along the way, rather than initially overbuilding and then encountering stiff resistance to a steep learning curve accompanied by constant time pressure and political scrutiny.

> This suggestion requires considerable forethought on the part of innovators and also requires that collaborations between innovators and the schools begin early in the process of innovation development and implementation. It is necessary to accommodate constraints and competing priorities of the system, to highlight some principles more than others, and to achieve some success, convince and persuade. Careful attention must be paid to the importance of trust and tenacity" (Blumenfeld et al. 2000, p. 159).

In summary, the authors wrote,

> The research and development community needs to develop its own capacity for generating the research base and the ability to work with the

> practitioner community to create useable reform programs and materials. (p. 162)

No approach to design is complete without careful attention to outcome measures and their influence on the process of learning. A systemic perspective recognizes the critical role of feedback in the process. Accurate feedback guides a system toward improvement; inaccurate feedback distracts or, worse, leads to chaos, poor decisions, and suspicion and political fractiousness. High stakes testing and accountability play the role of a brute force outcome measure, providing only the most rudimentary of feedback—"overall" improvement. With disaggregation of data by subgroups, a light is cast on inequity in many educational outcomes, but the high stakes testing data provide little guidance toward better solutions. Furthermore, the analytical tools for such outcomes have been rudimentary, not even supporting most users to look at combinations of categories of students.

The investment in Race to the Top and new assessments was founded in part on the hope of seeing advancements in the measurement of student progress. While some released items show promise, the structure of the planned reporting systems is coming up desperately short. Reports from high stakes tests represent an important measure of the overall system, but they provide little or no specific instructional guidance for teachers and students. The tests are too expensive for many districts, the data are likely to reveal disparities, and low performance on the more rigorous tests may exceed Americans' tolerance for bad news about their children and their schools. New approaches to testing must be designed and implemented.

The innovations and exciting developments that catalyze real change will not come from high stakes testing. End-of-year testing should be relegated to using student sampling just to get general measures of the well-being and progress of a school or district—similar to what NAEP assessments do at the state and national levels. Instead, the development of real diagnostic measures in STEM education would propel us forward. Interactive diagnostic assessment systems (IDAS) (Confrey et al. 2011) that provide a better means of gauging progress would offer several advantages in addressing these challenges and could incorporate mechanisms that support the instructional core for both teachers and students:

1. Data would be gathered through assessments that are embedded within technological environments, such as our formative assessments or measures of levels of participation.

2. The assessments would measure progress in student performance and understanding in real time.

3. Analysis of student progress would be based on growth, within a system of examination that allows the testing of conjectures (by teachers and by researchers) regarding why certain patterns in student ideas and responses emerge.

4. The assessment results would be tied to suggested interventions— targeted, tangible sources of help in addressing weaknesses. Using these would also allow schools themselves to see if they are functioning successfully as networked school improvement sites (https:// www. carnegiefoundation .org/our-ideas).
5. Finally, the assessments would be for learning, and therefore would be open to constant challenge by the students themselves to create a competency-based orientation to learning and growth.

Final Reflections

Systemic reform that serves urban children *must* place them, as next generation adults, at the center of the enterprise. When I directed the creation of the middle grades curriculum project (2012 to 2014), our team worked with urban children on a daily basis to try out materials and to learn what was needed to support their learning with the demanding curriculum (Confrey and Maloney 2015). The setting was alien to their experience. They typically lacked fundamental knowledge, and they had been accustomed to hiding in classrooms, mumbling, and forgetting what they had learned; our sessions with them shone a light on the effects of low expectations. The materials, however, were designed to engage and immerse them in mathematical practices by bringing the math to their own personal and sociopolitical worlds. Using problems oriented around everyday experiences or interests, such as texting, analyzing data on reaction time, exploring gerrymandering, and analyzing the make-up of the Senate, the students began engaging deeply in the mathematics, driven by curiosity and relevance of context. In addition, students were provided opportunities to be "doers" of mathematics, given an open workspace, and provided tools with which to work. As we watched and worked with them, we revised the materials, identifying key gaps in understanding and modifying the materials to address those gaps.

This approach capitalized on active learning, pioneering the concept of learning trajectories joined explicitly with the notion of productive struggle. The sequencing of the units was built from the concept of learning trajectories, as elaborated by Confrey and colleagues (Confrey et al. 2017, Maloney et al. 2014,), in which the challenges (tasks) were carefully sequenced to build solidly on the base of research on student learning. These sequences provide a fairly exact specification of the likely cognitive changes necessary to achieve an increasingly robust understanding of a big idea; they are based on empirical study, rather than on a logical analysis of disciplinary prerequisites. For example, in ratio reasoning, students move from an understanding of equivalence, to the specification of a base ratio, to a unit ratio/rate, and then to visualizing graphically, making comparisons, building up pairs of values, and eventually to a multiplicative relationship among the four quantities (Confrey et al., in progress).

Building from Kapur's (2008, 2010) notion of productive failure, students were asked to take on challenges that invited them to invent the idea as the beginnings of a solution to a problem and take time to work together during the launch phase. During the "dig-in," students were asked to create a possible solution to the challenge, such as predicting if and when a wind-up toy might fall off the table when set at a certain distance from the edge. They commented positively on the opportunity to craft the mathematics for themselves, rather than having it spoon-fed as a set of procedures unmotivated by a need.

During the final phase "share out," student strategies were exchanged and, within the context of interactive discourse, "formalizations" of the concepts were provided and briefly explored in the context of the challenge. For example, for any ratio relationship, there are two unit ratios. In a hot dog eating contest analysis, there are two unit ratios — one expressing how many hot dogs are eaten in a minute and one representing how many seconds it takes to eat an individual hot dog. Most students, and even the standards themselves, pay little attention to this duality in the concept, acting as if unit ratio/rate only involved a "per one unit of time." In my approach, the formalization is followed with an opportunity for students to make and experience the distinction between "number per unit of time" and "one unit of time per number of units."

I believe that, better than "productive failure" is to describe the process as *productive struggle*, because the final phase did support some students in obtaining the canonical solution and linking it to their struggles. However, as Kapur (2010) notes, other students will not do so: this is where the critical elements of differentiated instruction, scaffolding, and support must come into use. In our designs of tasks and instruction, we follow the challenges with a formative assessment to quickly provide teachers and students information on the degree to which the students had learned the critical idea(s) embedded in the task. In my original conceptualization, the challenges were to be followed by interactive demonstration, an active opportunity to provide additional opportunities for learning (but this part of the innovation was not realized in the published context).

Those middle grades curriculum materials also contained a set of math projects that provided students with opportunities to explore areas further, to synthesize what they learned, and to present it to each other in real and virtual performances. Again, these elements contributed to ways to encourage active and authentic learning experiences by children. Activating students through exceptional materials, as partners in the process of change is an essential component, but it is not sufficient to achieve the necessary goals as it lacks a true systemic framing.

In the beginning, I asked the question, "What was learned from systemic reform?" As one envisions a technologically rich future for schools, the lessons from systemic reform merit careful consideration. A commitment is necessary for **creating a *systemic* product** — one that pays equal attention to the materials and the relationships among the various participants and which characteristics were illustrated or foreshadowed in many of the systemic reform projects. The systemic reform initiative introduced many superintendents and principals to the idea of involvement in instruction and beyond regular administrative tasks. Systemic reform made clear the need to include attention to all partners, not only teachers and administrators but also parents, business partners, and informal science organizations in the communities. It required schools and districts to regard the combination of standards, curriculum, professional development, and formative classroom assessment as an integrated system. And it demanded **data be linked to the other factors to act as a form of feedback**. Not all of this was accomplished fully, but strides were made in the right direction.

Systemic reform has guided my own thinking on what the next generation of curricular products should include:

1. Clever ways to situate classroom activity in a journey shared by teachers and students through a set of standards, instead of using the tired notions of organizing instruction in units and lessons, day by day by day.
2. Deeply interactive systems in which students work actively in their own workspace and can exchange ideas and collaborate frequently, quickly, and easily, and in which teachers can actively orchestrate interactions.
3. Valid and reliable outcome measures of student progress linked to embedded assessments that gauge cognitive patterns and participation, with results provided to students, teachers, and parents.
4. The analysis of student progress with targeted interventions based on educational/cognitive needs.
5. Ways to facilitate the professional development and workplace behaviors of teachers and other professionals as members of a community dedicated to making change on behalf of students' futures.

These capabilities for a technological innovation can be built, and they can create the conditions for supporting many possible curricular materials, assuming each is carefully **tagged into a structured system of standards and learning progressions**.

If a core system had been structured as a learning system, dedicated toward self-examination of data early and often, and then released by a set of users responsible for protecting their own data and making their own adjustments and decisions after an initial period of implementation, then the innovation could express key lessons from systemic reform to create real change. Involvement with systemic reform, the ideas, the technical assistance teams, and most of all the determined practitioners frames much of my current work, as it sets a key standard for what success can look like.

References

Bill and Melinda Gates Foundation. (February 2010). Empowering effective teaching: Strategies for reform. https://docs.gatesfoundation.org/Documents/empowering-effective-teachers-empowering-strategy.pdf.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling up technology-embedded, project-based science in urban schools. Educational Psychologist 35(3), 139–64.

Chang, H. H., Honey, M., Light, D., Moeller, B., & Ross, N. (1998). The Union City story: Education reform and technology students' performance on standardized tests.

City, E. A., Elmore, R. F., Fiarman, S. E., & Teitel, L. (2009). Instructional Rounds in Education. Cambridge, MA: Harvard Education Press.

Cohen, D. K., & Ball, D. L. (2000). Instructional innovation: Reconsidering the story. Ann Arbor, MI. 1001, 48109.

Confrey, J., Gianopulos, G., McGowan, W., Shah, M., & Belcher, M. (2017). Scaffolding learner-centered curricular coherence using learning maps and diagnostic assessments designed around mathematics learning trajectories. ZDM — International Journal on Mathematics Education. In press.

Confrey, J., & Maloney, A. P. (2015). A design research study of a curriculum and diagnostic assessment system for a learning trajectory on equipartitioning. ZDM — International Journal on Mathematics Education 47(6), 919–32.

Confrey, J., Maloney, A. P., Nguyen, K. H., & Rupp, A. A. (2014). Equipartitioning, a foundation for rational number reasoning: Elucidation of a learning trajectory. In A. P. Maloney, J. Confrey, & K. H. Nguyen (eds.), Learning Trajectories in Mathematics Education, 31–60. Charlotte, NC: Information Age Publishing.

Confrey, J., Shah, M., McGowan, W., Gianopulos, G. G., & Jones, R. S. (in progress). Using a diagnostic tool based on learning trajectories to approach ratio reasoning.

Elmore, R. F. (2000). Building a New Structure for Educational Leadership. Washington, DC: Albert Shanker Institute.

Frye, M. (1983). The Politics of Reality. Berkeley, CA: Crossing Press.

Gray, L., Thomas, N., Lewis, L., & Tice, P. (2010). Educational Technology in U.S. Public Schools: Fall 2008. Washington, DC: National Center for Educational Statistics.

Honey, M., Culp, K. M., & Carrigg, F. (2000). Perspectives on technology and education research: Lessons from the past and present. Journal of Educational Computing Research, 23(1) 5–14.

Hug, B., Krajcik, J. S., & Marx, R. W. (2005). Using innovative learning technologies to promote learning and engagement in an urban science classroom. Urban Education, 40(4), 446–72.

Kapur, M. (2008). Productive failure. Cognition and Instruction 26(3), 379–425.

Kapur, M. (2010). Productive failure in mathematical problem solving. Instructional Science 38(6), 523–50.

Maloney, A. P., Confrey, J., & Nguyen, K. H. (eds.). (2014). Learning over time: Learning trajectories in mathematics education. Charlotte, NC: Information Age Publishers.

National Education Association. (2008). Technology in schools: The ongoing challenge of access, adequacy and equity. Washington, DC.

O'Day, J. A., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (ed.), Designing coherent education policy: Improving the system, pp. 250–312. San Francisco: Jossey-Bass.

Stokes, L. M., Sato, N. E., McLaughlin, M. W., & Talbert, J. E. (l997). Theory based reform and problems of change: Contexts that matter for teachers' learning and community (Final Report to the Mellon Foundation). Stanford, CA: Stanford University Press.

U. S. Department of Education. (2016). Improving basic programs operated by local educational agencies (Title I, Part A): Laws, regulation, and guidance. https://www2.ed.gov/programs/titleiparta/legislation.html

Weiss, I. R., Pasley, J. D., Smith, P. S., Benilower, E. R., & Heck, D. J. (2003). Looking inside the classroom: A study of K–12 mathematics and science education in the United States (Highlights Report). Chapel Hill, NC: Horizon Research.

Wenglinsky, H. (2000). Does it compute? The relationship between educational technology and student achievement in mathematics (Report). Princeton, NJ: Educational Testing Service.

Jere Confrey is the Joseph D. Moore University Distinguished Professor of Mathematics Education at North Carolina State University. She directs the SUDDS research group ("Scaling Up Digital Design Studies") in building new learning maps and related diagnostic assessments to support personalized learning, developing the application Math-Mapper. Confrey served as the Chief Mathematics Officer of Amplify Learning, 2012–14. She served on the National Validation Committee on the Common Core standards and built www.turnonccmath.net, a website that unpacks the Common Core. She was the vice chair of the Mathematics Sciences Education Board, National Academy of Sciences (NRC), 1998–2004; she chaired the NRC committee that produced *On Evaluating Curricular Effectiveness*; and she was a coauthor of NRC's *Scientific Research in Education*. She worked extensively on advisory boards and in technical assistance on systemic reform (SSI, USI, RSI). She co-founded the UTEACH program for Secondary Math and Science teacher preparation at the University of Texas at Austin. She authored Math Projects, Function Probe, Precalculus Interactive Diagrams, Graphs N Glyphs, and LPP-Sync software. Dr. Confrey received the PhD in mathematics education from Cornell University.

Chapter 8

Education, Equity, and Evaluation: From Reforming to Transforming Urban Systems

Bernice Taylor Anderson

The *2000 Biennial Report to the United States Congress*[1] by the Committee on Equal Opportunities in Science and Engineering stated that efforts to increase the flow of skilled U.S. workers must begin with the reform of K–12 education and that high quality education is a particularly relevant issue with regard to minority children. The report also called attention to the Urban Systemic Initiative (USI) program as one of the NSF-sponsored programs to address the shortcomings of American science and mathematics education:

> Urban Systemic Initiative (USI) program fosters partnerships between urban school districts and two- and four-year colleges and universities that conduct educational research. Projects are designed to:
>
> • Increase student achievement and enrollment.
>
> • Improve implementation of standards-based, inquiry-centered K–12 curricula.
>
> • Improve the competency and diversity of science and mathematics teachers in school districts that serve the largest number of school-aged children living in poverty.[2]

The 2011–12 report of the Committee on Equal Opportunities in Science and Engineering (CEOSE) restated the following alarm to the science and engineering communities:

> To retain the Nation's tradition of STEM leadership and to solve America'scompetitiveness dilemma, the country must increase its efforts toward successfully educating those groups within the Nation that are not being sufficiently tapped for the STEM workforce, particularly African Americans, American Indians, Hispanics, women, and persons with disabilities. … The more time that passes without action, the more urgent the problem becomes.[3]

Then CEOSE made the following recommendation:

> NSF should implement a bold new initiative, focused on broadening participation of underrepresented groups in STEM that emphasizes

> institutional transformation and system change; collects and makes accessible longitudinal data; defines clear benchmarks for success; upports the translation, replication and expansion of successful broadening participation efforts; and provides significant financial support to individuals who represent the very broadened participation that we seek.[4]

The committee further emphasized that "this initiative might include several multisite, geographically-based, national experiments … designed to significantly advance broadening participation across all levels of schooling, resulting in sustainable pathways pre-K–20+. … We see game-changers to reach the coordinated coherence and collaboration needed for a seamless pre-K–20+ scope."[5]

As the CEOSE recommendation is capturing the attention of senior NSF leaders and the various directorate advisory committees, those familiar with the systemic initiative era think of this call to action as the next generation of systemic reform. It is timely that as a bold approach to broadening participation becomes a priority topic for national conversation, NSF is also reviewing the foundation's decades of achievements in promoting diversity, equity, and accessibility in the science and engineering enterprise. Therefore, it is appropriate that this review of the lessons learned from the urban systemic portfolio contributes to future opportunities for transformative research and models or practices designed to "better serve groups traditionally underrepresented in STEM"[6] and to indicate how the emphasis on performance and evaluation has advanced diverse perspectives about documenting results and determining success. The following discussion highlights both the successes and the shortcomings of the USI/Urban Systemic Program (USP) in terms of scope, coordinated coherence and collaboration, talent development, and accountability.

Scope

The first attempt to transform whole systems in K–12 education began with the systemic reform approach, designed to build capacity to demonstrate that all students can learn and achieve at much higher levels in science and mathematics than were being achieved in the 1990s. It became the "bold experiment" that laid the infrastructure for systemwide models to improve collaborative performance among stakeholders to strengthen academic experiences and outcomes for all –12 students. Former NSF Director Neal Lane described the new investment as a "re-engineering" effort to support systemic approaches for far-reaching impact on improvements in the entire delivery system of precollege education. Under the leadership of Luther S. Williams, former assistant director of the NSF Directorate for Education and Human Resources, the USI/USP became the "revolutionary vehicle" that catalyzed change in the largest U.S. urban school systems with high concentrations of students in poverty and racial/ethnic minorities. In addition, to help sustain significant systemic improvements, the urban program included options to foster innovative partnerships between urban school districts and two- and four-year colleges and universities.

Systemic efforts increased the breadth and depth of NSF's impact on STEM education for over two decades. To varying degrees of implementation and accomplishments, USI fostered experimentation; accelerated the rate of policy changes; increased systemwide improvement in student learning in mathematics, science, and technology; and developed a high quality science and mathematics instructional and technological workforce. The USI program contextualized for urban schools the systemic lessons learned from the Statewide Systemic Initiative (SSI): "Education reform goes beyond the formal pre-K–12 system and must embrace institutions outside of the systems — e.g., families and community organizations, businesses, and institutions of higher education (IHE) — because student academic performance is affected by what occurs outside of school, not just in school; in other words, the 'system' being reformed consists of multiple institutions not just the pre-K–12 system."[7]

However, the system of multiple institutions was only held accountable for K–12 outcomes linked to aligned resources and mutual partnerships for improving K–12 preparation for all students. There was not an expectation of changing all participating institutions for a much broader diversity goal for the scientific enterprise; instead, the convergence of resources was designed for changing urban schools to ensure that the instructional/learning environment and support structures were improving precollege mathematics and science education. The coalition building among the partners tended to be focused primarily on the immediate goal of changing urban schools toward inquiry-based learning aligned with state and national standards that in turn did lead to educational improvements in all of the funded districts. (Other units in the Education and Human Resources Directorate were to assume the responsibilities related to changing higher education practices, the STEM workforce, and individual components of educational systems.) In addition, it is important to note that the size of the systemic reform investment required the NSF to play a catalytic role for change that was tailored to local urban needs within the context of the systemic reform drivers. The initiative was catalytic for strengthening the infrastructure for improving science and mathematics instruction across the precollege levels or elementary, middle, and high schools, rather than providing direct financial support to students to consider and/or pursue STEM careers. Other areas that were not directly addressed as part of the scope of urban education reform included (1) transforming all partners and (2) plans for sustainability were not fully researched or implemented due to the implementation challenge of going to scale.

These shortcomings can be linked to the duration of the program and size of the investment. Yet, each must be considered at the outset of the next large-scale, full scope investment in mathematics and science education designed to tap the talent of underrepresented populations. A future call for a pre-K–20+ effort for broadening participation might be even more successful if such an initiative supported high-performing partnerships for approximately 20 years to move the needle and to be able to measure success using cross-sectional and longitudinal analyses. Moreover, the earlier experiments with systemic reform have prepared the nation to be innovative and expansive such that there can be a call for a large-scale, comprehensive initiative that links multiple institutions and connects all relevant stakeholders for transforming the scientific enterprise by ensuring the full participation and advancement of underrepresented groups in STEM learning and research across multiple venues nationally and internationally. Once again, any new initiative needs to be designed for sustainability, remembering the following observation one of the evaluation firms that conducted systemic reform project-level evaluations said: "There is a slow, steady stream of benefits that continuously accrue in a district when an improvement infrastructure is maintained."[8]

Coordinated Coherence and Collaboration

The USP evaluation by COSMOS[9] stated that genuine reform required districts to demonstrate commitment by enacting a comprehensive vision that emphasized the mechanism(s) of change. It was important for districts to focus on the critical alignment of key components of the K–12 system, but it was far more essential to build continuous momentum in support of all students reaching their full potential and receiving the needed support from many stakeholders who embraced and aligned efforts and innovative ideas around the common theme of high quality science for all. In other words, partners were to envision and advance the attainment of a reformed education system through the convergence of resources where standards-based curriculum was aligned with instruction and assessment and a coherent set of STEM policies supported reform goals with broad-based support from parents, policy makers, administrators, teachers, and collaborators from higher education, business and industry, and informal science communities. This vision would be realized when there was a broad and deep array of evidence of enhanced student achievement and improvement in the achievement of all students.

USP sites aimed to close or reduce the achievement gap by providing high quality, standards-based instruction to all students. This inclusive agenda did have promising outcomes as all districts documented positive pre- and post changes on a range of assessment instruments. Closing the gap required so much focused attention and strategic use of resources that the urban systemic efforts did not directly have as an explicit goal the longer-term goal of increasing the number of students who pursue STEM careers. The coordination effort with higher education institutions was for professional development opportunities for K–12 teachers. In later years, some of the sites were able to receive additional resources for teacher preparation models and technician education. The lesson learned, however, was that future multiple level partnerships should involve shared processes and outcomes for precollege and postsecondary levels of the educational continuum. This is the next generation of reform efforts that is being advocated by the CEOSE.

With urban systemic reform, there was coordination across multiple components simultaneously. As a result, there was a limited opportunity for urban school districts to collaborate in the conduct of foundational education research. The few that did engage in such became the testbeds for research on practice. More research on practice as well as theoretical research is needed to guide such change. In addition, with the current focus on transparency, future efforts can learn from USP that successful coordination and visioning as well as a long-term commitment of public resources do and will require mobilizing public support that includes community collaboration and meaningful engagement in both the education and research components of a “bold new initiative.” In the systemic change environment for discovery and/or learning, support to large systems needs to be conceptualized, designed, and evaluated as “value-added investments.”[10]

Talent Development

“According to data collected from 1998–2002, close to 13 million students, more than 600,000 teachers and almost 51,000 administrators (principals, science, mathematics and curriculum leaders, etc.) were supported by NSF funds.”[11] These numbers continued to increase as the USP ended its funding to the last cohort of urban schools in 2006. It is important to point out that evaluation results did indicate better prepared students in mathematics and science. Relatedly, the USP focus on increased student achievement was accompanied by higher expectations for increasing the enrollment of traditionally underrepresented groups in higher-level mathematics and science courses. The accomplishments and demonstrated impacts did not include student outcomes beyond the high school years.

Former NSF Director Rita Colwell applauded the program for showing that all students, regardless of their backgrounds or surroundings, can tackle challenging mathematics and science courses when school systems use investments wisely to support systemwide policies for learning, to develop capabilities of teachers, and to connect with the community through partnerships.[12] The collaboration with four-year colleges and universities was impactful with early career educators in the area of professional development. The heavy focus on professional development resulted in a major investment in the STEM instructional workforce to ensure high quality precollege STEM education. “An average of 49% of USI funds were spent on professional development activities across all of the USI sites”[13] to improve teachers’ content knowledge, pedagogical skills for inquiry-based teaching, and methods of assessing student learning. Contributing to better classroom instruction and better prepared students to attend college by focusing on better equipped teachers was a successful strategy employed by sites. However, boosting college enrollment was not tracked consistently to document or to determine the USI/USP highly plausible contributions to increasing the number of high school graduates from underrepresented groups who considered and/or graduated with STEM degrees and/or certificates.

Accountability

A high-stakes accountability-driven management approach for implementation and measuring results characterized the systemic reform movement. The drivers framed the context for coherence, collaboration, and assessment. The progress across sites was uneven and with any new initiative in the future, the success of the efforts can be leveraged and evidenced when the awardee has a very ambitious, comprehensive plan designed to make significant impacts and with strong quality control mechanisms. For example, the leadership in the Jacksonville USP pointed out that "all the work we do is grounded in data-driven research."[14] The urban sites employed new practices where data were used for establishing goals, guiding instructional planning, and informing strategic decision making.

The complexity of system change required a different approach for assessing progress, measuring results, and reporting outcomes. Major improvements in identifying and/or developing credible data sources took place to address a range of evaluation needs. The districts increased their evaluation capacity by supporting internal and external evaluators, and NSF expanded its evaluation framework, emphasized mixed method designs, and employed multiple external studies for evaluation and research purposes. An integrated system of (1) monitoring annual performance at the project and program levels on common measures; (2) conducting regularly scheduled internal and external expert reviews, including programmatic reviews by the Committee of Visitors, performance monitoring reviews by the EHR Senior Leadership Team, and midpoint reviews of cohorts of awardees; (3) assessing project-level evaluation findings included in the required NSF annual project reports and responsiveness to qualitative reporting guidance; (4) sharing and synthesizing a set of competitively funded research and evaluative studies; and (5) using the findings and recommendations of a third-party program evaluation had to be established and successfully implemented to meet accountability demands and promote the foundation's growing interest in evaluation for learning. The emphasis on accountability expanded the evaluation knowledge base about technical or methodological challenges, as well as the cultural, political, and practical issues about evaluating systemic reform. In assessing a complex experimental effort, a key lesson is to call for nontraditional methodologies and analytical tools to document and evaluate success, as well as using negative findings to promote improvement. Future large-scale initiatives need to require the use of appropriate, mixed-method designs from multiple sources of information and data in which a systemic or system change evaluation considers different disciplinary approaches to define and measure success.

In closing, the program evaluation of USP indicated that systemic reform was driven by multiple processes that have their own pace and force and that there can be reversal due to other external conditions, such as financial and policy factors. Surely this has happened since the urban systemic program evolved into other less comprehensive reform initiatives due to federal changes in perspectives, resources, and approaches to solving the grand challenges in STEM education. As urban districts seemed to have lost some of the collective strength for a coordinated federal approach for bottom-up change, it is timely and deemed appropriate for districts and their surrounding communities to become reenergized and take hold of the new CEOSE 2011–2012 recommendation for the STEM enterprise. Pursue copartnering with higher education, business and industry, professional associations/societies, informal science organizations, and so on, to reinvigorate a transformative STEM education delivery system that includes underrepresented populations across levels and within all partnering organizations for significant individual, institutional, disciplinary, and interdisciplinary outcomes. Leading a new generation of multi-sector partnerships will require innovation in both broadening participation implementation research and broadening participation evaluation studies. It is time for the urban community to reclaim its innovative role in championing diversity and inclusion in STEM education and to demonstrate renewed or new capabilities in leveraging assessment and evaluation data for developing new science and engineering paradigms for broader participation.

Moreover, with the expansion of partners, the notion of going to scale will have to encompass new perspectives. For example, how can there be scaling from district-wide reform of schools and communities to reaching regional empowerment of underrepresented groups for economic competitiveness to enabling sustainable models for national leadership on inclusion in the science and engineering workforce? The urban systemic reform movement initiated in the 1990s was a grand experiment that introduced positive changes to the K–12 system only; now it is time to leverage the urban power for a new reform framework that addresses a seamless STEM education system driven by engineering the "alignment" of innovative solutions that address the issues of diversity, equity, accessibility, performance, advancement, and empowerment.[15] The timing is crucial and, since there is momentum among various groups in support of another bold game changer in STEM, it is essential that the leaders of urban systemic reform revitalize the urban community to have a highly visible presence in shaping the transformative and system change response to broaden participation in science and engineering.

References

Committee on Equal Opportunities in Science and Engineering. (2013). 2011–2012 Biennial Report to Congress: Broadening Participation in America's STEM Workforce. Arlington, VA.

Committee on Equal Opportunities in Science and Engineering. (2000). 2000 Biennial Report to the United States Congress. Arlington, VA.

Committee on STEM Education of National Science and Technology Council. (May 2013).

Federal Science, Technology, Engineering, and Mathematics (STEM) 5-Year Strategic Plan. Washington, DC.

COSMOS Corporation. (2005). Cross-Site Evaluation of the Urban Systemic Program — Final Report: Strategies and Trends in Urban Education Reform. Bethesda, MD.

COSMOS Corporation. (n.d.). Findings Brochure for the Study of Statewide Systemic Reform in Science and Mathematics Education. Bethesda, MD.

Inverness Research Associates. (2007). Lessons Learned from the San Diego Urban System Project (USP): Implication for Funders and Future Project Designers. Inverness, CA. ERIC ED 5000552.

Inverness Research Associates (2006). The Portland Urban Systemic Program (USP) — Five Years of Building Systemic Support for Math and Science Education Improvement: 2001—2006. Inverness, CA.

National Science Foundation. (2001). NSF Press Release 01-53: Big City Students Make Gains in Math and Science, Report Says. Arlington, VA.

National Science Foundation. (2000). Urban Systemic Program in Science, Mathematics and Technology Education: A Foundation for K–12 Science and Mathematics Educational System Reform. NSF 00-34. Arlington, VA.

Potomac Communication Group. (2005). What Works in Science and Mathematics Education

Reform: A Report of the National Science Foundation's Urban Systemic Program. Washington, DC.

Systemic Research, Inc. (2001). Academic Excellence for All Urban Students, Norwood, MA.

[1] Committee on Equal Opportunities in Science and Engineering. 2000 Biennial Report to the United States Congress.

[2] Ibid., p. 13.

[3] Committee on Equal Opportunities in Science and Engineering. 2011–2012 Biennial Report to Congress: Broadening Participation in America's STEM Workforce. July 2013, p. iii.

[4] Ibid., p. 22.

[5] Ibid., pp. 21–22.

[6] Refer to CoSTEM report where this is a goal for all federal agencies. Committee on STEM Education of National Science and Technology Council. Federal Science, Technology, Engineering, and Mathematics (STEM) 5-Year Strategic Plan. May 2013.

[7] COSMOS's findings brochure for the Study of Statewide Systemic Reform in Science and Mathematics Education, n.d.

[8] Concluding statement in the Inverness Research Associates' evaluation of the Portland USP 2001–2006, p. 33.

[9] COSMOS Corporation. Cross-Site Evaluation of the Urban Systemic Program — Final Report: Strategies and Trends in Urban Education Reform, 2005.

[10] Researchers at Inverness Research Associates conceptualized the San Diego USP grant as a value-added investment, adding value to existing efforts, and did not expect the funded work to single-handedly produce unrealistic outcomes.

[11] Potomac Communication Group, What Works in Science and Mathematics Education Reform: A Report of the National Science Foundation's Urban Systemic Program, 2005, p. 1.

[12] NSF Press Release 01-53. Big City Students Make Gains in Math and Science, Report Says. June 28, 2001.

[13] Systemic Research, Inc., Academic Excellence for All Urban Students, 2001.

[14] Potomac Communication Group. What Works in Science and Mathematics Education Reform.

[15] Note that galvanizing the community is also applicable to the rural systemic initiative.

Bernice Taylor Anderson is a senior advisor in the Office of Integrative Activities (OIA) at the National Science Foundation (NSF). She is also the executive secretary of the Committee on Equal Opportunities in Science and Engineering (CEOSE). She began her career at NSF as a program director in the Division of Research, Evaluation, and Communication in the Directorate for Education and Human Resources (EHR) in 1997 and advanced to several leadership positions in the EHR, including, senior advisor for Evaluation and Information Management, acting director of the Division of Human Resource Development, and acting director of the Division of Educational System Reform. Prior to joining NSF, she was a research scientist in the Special Populations Group in the Education Policy Research Division at Educational Testing Service in Princeton, New Jersey. She began her professional career as a faculty member in the School of Education at Norfolk State University, Norfolk, Virginia. She is the coauthor of the book *Breaking the Barriers: Helping Women and Minority Students Succeed in Mathematics and Science* and a book chapter, "Evaluating Systemic Reform: Evaluation Needs, Practices, and Challenges" in *Evaluation of Science and Technology Education at the Dawn of a New Millennium*. She received the Doctorate of Education degree from Rutgers University in 1984.

Chapter 9

The Retrospective and Reflections/Perspectives in the Aggregate: Lessons Learned and Recommendations

Luther S. Williams and Margaret (Midge) Cozzens

The systemic initiative (SI) programs were not undertaken absent an appreciation of the magnitude of the challenge and the attendant risk. In an effort to lower the risk, a series of pre-award actions and award implementation, and evaluation and accountability processes were elaborated and employed. These included (a) planning grants awarded to ensure a full understanding of the requirements and ability to comply; (b) cooperative agreements for each grant under which the funder and performer(s) regularly interacted; (c) major SI program elements or requirements, termed *process and outcome drivers*, to promote unitary addresses of the program variables, especially the productive synergies at the interface of the various system components; (d) early phase learning, with occasional redesigns of the ongoing system reform process, research, and evaluation of various types as described in previous chapters; (e) regular meetings of program superintendents and program directors and all SI program leaders; and (f) technical assistance. Technical assistance proved to be invaluable in facilitating the sharing of promising practices across sites.

A number of lessons can be learned from the systemic reform programs of the 1990s and early 2000, as we indicated in the previous chapters of this book. We reiterate those lessons here to emphasize their collective importance and to provide recommendations for the future.

1. **All students.** One should embed equity into the core system design and conduct of any K–12 program and then inform the resultant student learning outcomes with experiential knowledge, as contrasted with a great collection of a priori assumptions. It should be observed that such insistence does not presume that there are not some students who possessed limited facility to learn, just that all students must have the opportunity to learn at the highest level possible. Differential K–12 educational sequences existed in some of the school systems participating in systemic reform, and the educational justifications often appeared to negate the opportunity for some students to understand or express their talents and to develop the attributes required to

become productive learners. Regarding this matter, an omission of the SI programs was inattention to how the dynamics of microaggressions and marginality (Sue 2010) rendered the participation by and productive engagement in mathematics and science learning by some of the students in question even more problematic. This is not to suggest that these students are expected to counter such massive and deeply ingrained sociopolitical constructs, but it is reasonable to assume that they should have been instructed in strategies by which to mitigate some of the adverse impacts of such marginalization. To frame the dimensions of this in a larger context, in 1817, the Vermont Gazette wrote, “Let your motto be—internal vigilance is the price we pay for liberty” (Bennington 1817). Almost two hundred years later, we need to realize that the same is true for equity. Without explicit attention, equity moves from being a core component of high quality mathematics and science education back to being a special program or an “add-on.” As we move forward, it is key to acknowledge that high quality education is not only that which happens in the classroom. Disciplinary practices that lead to different punishments for similar behaviors and selection policies that lead to overrepresentation of some groups in special education and gifted education all need to be addressed if equity is at the core of education. The disaggregation of data continues to be key but only if those data are used with diagnostic tools and corresponding data-informed actions are taken. The standard to which we need to aspire remains. The longer-term goal is to serve all students so that outcomes can no longer be predicted based on sex, ethnicity, class, disability, or language group. “Unless all students are achieving in standards-based mathematics and sciences, no claim for high-quality education can be made”(Campbell and Jolly, Chapter 3, 2017).

2. **Teachers and their exemplary professional practice are paramount factors to student success.** Shortchanging teacher development in any reform or initiative undermines everything else. System change requires system strategies and system intelligence, but that requires resourced attention to the human interface among the entire society, schoolchildren, and the teacher! There will be a temptation in the digital era to discount the role of the teacher in the light of the other knowledge resources available to students. We are on the cusp of an era where many of the onerous and repetitive tasks of teachers can be offloaded, opening the opportunity for teachers to function at a much higher level than their position descriptions permitted in the past. The systemic reform efforts, canvassed in this volume, underscored in the 1990s how critical teacher development was. The current success of countries like Finland and Japan, which have operationalized a deep cultural respect for teachers and their professionalism, reinforce this truth! To be sure, important new developments in fields such as student data analytics and social and communication technologies are going to vie with exciting knowledge breakthroughs, in terms of competing for the attention of educational policy makers, reformers,

researchers, and other stakeholders in shaping next-generation learning and education. Yet, in the midst of exciting change, profound challenges, and breakthrough opportunities, effective teaching will always be the most important malleable variable in system's success.

3. **Curriculum, assessment, and teacher development are a required closely linked trio.** Setting clear and ambitious mathematics and science standards was shown to be an imperative to sustained improvements in student learning outcomes and the orderly implementation in the classroom of proscribed critical systemic components. We learned that standards-based instruction is not just a slogan, but that high quality curricular materials relative to the standards are essential and that teachers need to be prepared to use these materials. We learned that assessments are critical for the public view of education and that equity of educational opportunities is too often motivated by high stakes state and national assessments. But we also learned that assessments independent of curricular materials or teacher education and professional development won't work. In addition, we learned that increasing the mathematical or scientific content knowledge of teachers independent of the availability of quality curricular materials doesn't work and that this content knowledge needs to be related to what the teachers teach. We need not worry if teachers and students are learning at the same time; when that learning applies to the use of technology and in cross-disciplinary contexts, it is actually a plus. Serious attention must be paid to all the component parts needed for implementation of these standards, coupled with a reasonable set of rewards for teachers engaged in the reform effort.

4. **Reform capable infrastructure models for success.** Another powerful lesson derived from the implementation and progress of the various SI programs is that a robust set of analytical models must be developed to help guide the policy and resource-allocation changes that need to be made. The models should be a hybrid of system-dynamics–based and agent-based models that allow one to combine the interaction of system components with the impact of the likely behavior of the system's actors: teachers, students, principals, administrators, and community stakeholders. It is a daunting effort — actually, plainly impossible — for the system's leaders to intuit the complex, dynamic behavior of the system's multiple components, as policy and resource allocations are changing without the use of models. For instance, small-scale models can help determine whether a policy that undergirds teacher professional development workshops will foster a change in instructional style can work. A model of this policy to determine whether the goal of transforming the teacher workforce can be reached can integrate variables such as the effectiveness of the workshop training with the length of the workshop and the propensity of the teachers to actually participate based on workshop content and length. A larger scale model can foster understanding of the kinetics of system change by elucidating the interactions of major components

— professional development, curriculum, testing, teacher quality, and so forth — and the impact of these interactions on the overall goal of student learning. Such a model helps identify system leverage points that may be available to further and/or accelerate effective and sustainable system reform.

5. **The key actors in the reform process must be familiar with the methods and tools for fostering innovation and must be committed to making the innovation work**. Leadership is important. This is deemed important because many of the changes fostered by the reforms are innovations, and innovations are usually met with considerable resistance. Such proved to be the case in the instances of the SI reform programs addressed in this volume, and history has provided numerous other examples. Succinctly stated, system reform cannot simply be mandated; it must be owned by the system's actors, and, thus, innovations and the attendant values must be rendered convincingly evident, or sold! In addition, SI programs indicated the significance of developing an internal framework for the execution of a complex system mindset. Equally important is the understanding that orderly progression of such a complex system to a new and improved system requires sustained innovations but only if such innovations are configured in a multistep process. In the final analysis, nothing lasts unless it is supported in the budget of the system.
6. **Accountability from beginning to end and beyond is required.** The systemic initiatives required a plan of action before they were to be funded, and this plan of action focused on key process drivers and outcome drivers, as discussed in a number of the chapters in this book. Implementation of systemic activities from choice of curricula to professional development to funding coherence, and more, was monitored on a yearly basis — with site visits (both on-site and reverse site) and program officer interactions — by technical assistance providers. To link effect to cause in education is indeed difficult as we are not running controlled experiments on our children's lives. Yet we need tools to recognize when we are making a difference and when we are not, and to abandon those things that on the surface looked promising but fail to fulfill the promise. Thus, applied research that occasions the generation of new-age and diagnostically robust tools to deeply assess the utility of systemic elements in practice will guide overall program evaluation and modifications today and in the future.

Succinctly stated, the most important lessons of this volume are simply too important not to be learned. While the various programs differed somewhat in design and scale, common lessons learned that can be termed imperatives for reasonable success include (a) full or nearly complete implementation of a standards-based curriculum-instruction-assessment continuum, (b) a major investment in teacher quality, and (c) the fostering of systemic reform with a unitary approach and a coherent and sustained policy construct. In addition, it takes considerable time and effort to achieve a shared vision for reform; agreements on a reform strategy; adoption of a systemic approach; and a complex balancing act, one that involves simultaneous attention to multiple tradeoffs; and extensive evaluation.

Much has changed and will continue to change since the systemic initiatives of the 1990s, especially the innovations in learning sciences and learning technologies that alter the sense and experience of teacher professionalism, and the hope that these will be catalytic precursors to more sustained and systems-oriented changes in education. We have since learned some small things, but things that affect the whole system, such as, "When students and teachers learn together, everyone learns," and that part of any education, especially in mathematics and science, is to help develop habits of mind that last a lifetime. Children (those in urban, rural, or suburban settings) are our next generation adults and as such must be the center of any educational reform activities. Last, we need research to provide more sophisticated tools to understand the functioning and health of our educational system.

References

(Bennington) Vermont Gazette, July 8, 1817, p. 2. https://www.monticello.org/site/Jefferson/eternal-viligance — price-libertyquotation#footnote2 rmy50k3.

Campbell, P. B., and Jolly, E. J. (2016). Infusing Equity in Systemic Reform: Influencing the Understanding of Equity and Excellence, an essay in this series.

Reflections from Wingspread. (1994). "Lessons Learned About the National Science Foundation's Statewide Systemic Initiative: A Report of the Wingspread Conference." Chapel Hill, NC: Horizon Research.

Sue, D. W. (2010). Microaggressions and Marginality: Manifestation, Dynamics and Impact. Hoboken, NJ: Wiley.