Systemic Educational Reform in a Large Complex Educational System: Lessons Learned

Presented at the AAAS-UNESCO International Conference on Science and Technology Education: Systemic Approaches to Reform

By:
Manuel Gómez, Director
Resource Center for Science and Engineering
University of Puerto Rico
June 6-9, 2004
Paris, France
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Overview</td>
<td>1</td>
</tr>
<tr>
<td>3. The Essential Difference Between Science and Mathematics (S&amp;M) Reform Projects and Systemic Reform Process</td>
<td>16</td>
</tr>
<tr>
<td>4. Establishing a Quasi-causal and Persuasive Relationship Between the Systemic Reform Initiatives and the Improvement of Student Learning: The Use of Assessment and Attribution to Drive the Educational Reform</td>
<td>19</td>
</tr>
<tr>
<td>5. Validating the Transferability of a Systemic Reform Model</td>
<td>26</td>
</tr>
</tbody>
</table>

-i-
Overview

Realizing that previous isolated efforts to raise standards of teaching/learning and improve student achievement in the key disciplines of Science and Mathematics in a large educational system were ineffective, Puerto Rico began the systemic reform of Science and Mathematics education with National Science Foundation funds through the Statewide Systemic Initiative. Since educational systems are complex social systems that include multiple centers or elements of power with considerable autonomy and resources at their disposal, it is essential to come to terms with the reality that an institutional cultural transformation is required in order to put in place a fundamental reform that will affect Science and Mathematics teaching/learning, and that can be implemented and will endure.

The empirical theory of educational systemic reform, extracted from the lessons learned in this experiment, is a synthesis of the collective wisdom accumulated during more than a decade of achieving systemic reform that demonstrated that ALL students can learn Science and Mathematics if the proper environment is created. In 1993 when the reform was initiated, Puerto Rico’s public education system (after New York, the largest educational system in the United States) was a highly centralized, dysfunctional system, serving over 650,000 students, 80% of whom come from families below the federally established poverty level, in 1,500 schools with 30,000 teachers. At the end of 10 years, the reform had effectively reached 800 schools, more than 50% of the system’s schools.

It became evident that if reform was to happen, the first challenge was to understand the complex educational system as well as how to influence its behavior. Complex social systems are highly dynamic and reactive (see paper in Appendix). In contrast to physical systems, complex social systems’ web of interactions and feedback are subject to constant modification in response to internal and external stimuli. They are also designed to sustain the status quo and have a natural tendency to seek stasis. Such social systems are very sensitive to external evaluation and assessment and react by modifying their behavior only to the extent needed to accommodate the external challenge. They are organized to preserve their institutional culture. A complicating factor for education reform is that the educational system is strongly coupled to the body politics and public opinion that control resources and infrastructure and can impose constraints through legislation and monetary disbursements. All of this reinforced our belief that, before taking action, one factor was of utmost importance: fully understanding the nature of the system’s interactions and feedback.

Lessons Learned
First: success depended on a multi-pronged approach that would affect the system at key places where optimal leveraging could be achieved. Six key elements of the system must be impacted simultaneously:

- The policy-making level at the district and at the Department of Education level
- The curriculum and educational standards widely accepted by system constituencies
- The empowerment of teachers, principals, and schools to change the teaching/learning environment.
- The school was taken to be the unit of change, and the transformation of the school into a Community of Learners was found to be the most effective strategy to transform the school’s teaching/learning culture.
- Universities’ Schools of Education and Natural Sciences, the source of future teachers needed to become active participants in the reform of teacher preparation programs and actively participate in in-service teacher preparation programs, following a train-the-trainer approach.
- The body politics and the community needed to be actively involved to ensure understanding, support, and sustainability of reform.

Second: top leadership--the Governor, the Secretary of Education, Superintendents, and the Legislature--must simultaneously put in place policies that support reform efforts. In Puerto Rico, key legislation (Laws 18 and 149) provide specific mechanisms for the implementation of educational reform by requiring all schools in the system to become community schools (site-based management), vesting them with academic, administrative, and fiscal autonomy while remaining accountable to the central system for the achievement of high standards of excellence. These laws also mandated the decentralization of the educational system; redefined the role of the central and regional officials to that of facilitators; promoted the active participation of teachers, parents, and the community in the educational process through formation of School Councils; and created a career ladder that allowed teachers to be promoted on the basis of their teaching performance and their students’ academic achievement, thus serving as a self-correcting academic accountability process.

Third: the reform effort could work in the long run only if the existing educational system resources are harnessed and used optimally and effectively to scale-up reform and achieve new effective and enduring ways of educating students. This proved to be one of the biggest challenges.
Fourth: start small and scale-up. This permits the careful monitoring efforts, using data to establish a “quasi-causal” relationship between reform activity and improved student performance to provide for timely corrections and ensure quality control in the scaling-up process.

Fifth: a self-correcting feedback loop was developed to allow the system to correct itself through feedback by bringing together assessment, attribution, and accountability to interact with each other: a process we called A³ for short.

**The Need for a Unit of Change and a Scale-Up Strategy**

The Puerto Rico educational system’s size and complexity and the limited resources to achieve reform meant the whole system could not be tackled initially: hence the need for a unit of change. The reform chose the school as the unit of change, a choice that was made possible by the simultaneous enactment of the Community School Law that provides administrative, academic, and fiscal autonomy to the school. The teaching/learning and institutional cultural transformations took place at the school level, using a “Whole School Approach” that required all science and mathematics teachers and the principal in the school to recognize and accept the need for reform, and, with the appropriate training, evolve into a Community of Learners to carry on reform beyond the initial professional development stage; thus resulting in the school’s cultural transformation that provided a nurturing environment for teachers’ and students’ enhanced learning.

Dissemination Centers created in each educational region became the strategy for scaling-up the reform. The Centers were schools whose staff and teachers were imbued in the reform principle; their teachers became leaders in providing technical and academic assistance, building collaborations, and offering professional development to schools in their region. Laws 18 and 149 institutionalized professional development and accountability systems as well as the use of assessment supported systemic reform and motivated other schools to join the scaling-up process.

**Value-Added by Systemic Reform**

Results of pre and post-tests developed to assess the effectiveness of the reform showed that reform schools and students performed better in mathematics and science open-ended questions measuring depth of understanding and higher order thinking skills than they did in multiple-choice questions based more on memorization of facts. According to Dr. Edward Kifer’s research, Japan is the only country that has shown similar student achievement patterns.

The results of the College Entrance Examination Board tests, the equivalent of the SAT test in the United States, showed a 50% reduction in the gap in performance of students from fully implemented-reform public schools and students from private schools, which
are essentially middle and upper class student schools. The significance of these results is that low-income family students can now successfully compete for admission to Puerto Rico’s most competitive universities.

After 10 years, the impact of reform has been sustained at the unit of change, the school, despite the fact that many of the policies have reverted due to current political realities. This demonstrates that the fundamental changes made in transforming schools’ teaching and learning culture and the system’s wide approach to change institutional culture endures in and above political variability.
An Empirical Theory to Guide Systemic Reform

NSF started the systemic reform effort upon realizing that, in order to raise standards of teaching/learning and improve student achievement in the key disciplines of Science and Mathematics, it could not continue to sponsor only isolated efforts within educational systems. If fundamental reform in the way that Science and Mathematics teaching/learning was to take place, NSF needed to come to terms with the reality that what was needed was an institutional cultural transformation of a complex social system that included multiple centers or elements of power with considerable autonomy and resources at their disposal. The four major elements of the complex system are: the schools, districts, and States Departments of Education; the universities; the body politics; and the public.

NSF started its systemic reform project at a time when little was known about how to achieve the goal of systemic reform in the educational enterprise. After more than eight years of experience in this effort, NSF and the systemic projects have jointly walked up the learning curve and have accumulated enough experience and know-how to articulate an empirical theory of systemic reform for the educational continuum. The practitioners have learned much about how to modify the behavior of a complex and dynamic social system to achieve the common goal of providing standards-based education for all students that results in the improvement of their learning and performance in Science and Mathematics. The empirical theory presented here is a synthesis of the collective wisdom accumulated during more than eight years of experimentation, trial and error, and unquestionable success, of the eight remaining statewide systemic initiatives in achieving systemic reform.

A. Understanding Complex and Dynamic Social Systems

If systemic reform is to happen, the first challenge of the reformer is to understand the complex educational system and how to influence its behavior. Although much is known about the behavior and design of complex physical and biological systems, the body of knowledge on how to influence and modify the behavior of complex social systems is much more scant. An exception to this is found in health care reform where some key principles of social systems behavior have been put to good use.

Complex social systems are highly dynamic and reactive. They possess a complex web of interactions and feedback among their elements and, contrary to physical systems, the rules of interaction and feedback are subject to constant modification in response to internal and external stimuli. These complex webs of social interactions are designed to sustain the status quo and the natural tendency of social systems to seek stasis; they are very sensitive to external evaluation and
assessment and react by modifying their behavior by accommodating to the external challenge while preserving their institutional culture.

From the student point of view, the educational system is a continuum: the so-called K-16+ continuum. For the student's learning process, it is essential that all the elements of the system are aligned following a standards-based curriculum that is sequenced and coordinated to meet their specific learning needs throughout the K-16+ continuum and that well-defined educational goals exist at all levels. Unfortunately, the organizational structure of the continuum is typically fragmented into autonomous or semiautonomous units: the universities, departments of education, school districts, and schools. Furthermore, this educational system is strongly coupled to the body politics and public opinion which have control of resources and infrastructure and can impose constraints on the system through legislation and disbursements. As a result, the interactions and feedback between the elements of the system are less than perfect and usually result in a dysfunctional educational system that does not provide the student with an articulated educational continuum. Any systemic reform theory, and its resulting praxis, must address this issue; the practitioner of systemic reform must fully understand the nature of the interactions and feedback before acting.

The complexity of the interactions and feedback within the K-16+ continuum requires that the reformer follows a multi-prong approach to affect the system at key places where optimal leveraging can be achieved. In order to be systemic, five key elements of the system must be impacted simultaneously:

- The policy-making level at district and states Departments of Education must be affected.
- National educational standards must be put in place and widely accepted by the system’s constituencies.
- Teachers, principals, and schools must receive professional development that is aligned with standards and they must become empowered to change the teaching/learning environment of their schools.
- Universities' Schools of Education and Natural Sciences, the source of future teachers, must be actively involved in reforming their teacher preparation programs and stimulated to forge alliances with the K-12 educational system.
- Advocacy and effective communication must be established with the political domain and the community to ensure understanding, support, and sustainability of the reform.

Social systems are highly dynamic and non-linear in nature. Thus, sequential linear implementation of strategies is not appropriate because it fails to take these facts into account. A holistic approach is necessary where the synergy that characterizes
complex systems is strategically used to achieve the desired result. The time delay between the introduction of a strategy or activity and the time the system takes to respond must be, at least, intuitively understood by the reformer; otherwise well thought strategies can produce undesirable results. A sense of timing in the introduction of strategies and activities is of the essence; multiple strategies must be carefully orchestrated so that they form a coherent and harmonious whole. Multiple dynamic models for simpler systems exist; they can provide guidance and understanding of complex system dynamics and help to avoid the cancellation, or worst, counterproductive effects of poorly orchestrated reform models.

The behaviors of complex social systems are hard to predict in advance. Thus, the reformers must possess entrepreneurial attitudes and skills; they must have the willingness to take risks and to modify the implementation plan in order to take advantage of unanticipated windows of opportunity that arise as the system evolves.

B. The Reforms’ Resources are Never Commensurate with the Educational Systems’ Resources

The reform effort must be catalytic in nature because, in general, education systems have large resources, both fiscal and human, at their disposal which are committed to procedures and management structures that follow well established and proven paradigms. In contrast, reformers typically have available limited resources to modify the behavior of the system in order to achieve reform.

Because of the principle of stasis, educational systems have a natural tendency to resist change. When confronted with mounting evidence that the system is not performing effectively and external pressures for change increase, systems tend to effect changes that accommodate to their existent organizational and management structures rather than go into a “paradigm shift mode”. As a result, the system tends to become a collage of many accommodations that are internally inconsistent and result in dysfunctionality. The challenge of the reformer is to catalyze a paradigm shift in the educational system that will transform the system from a dysfunctional to a functional one while, at the same time, introducing institutional learning elements that will make the system self-correcting.

Given the large discrepancy in the magnitude of system resources and reform resources, strategic thinking is needed. Multiple foci of reform must be identified and developed both horizontally across the many elements of the system as well as vertically along the educational pipeline and the policy making levels for both the K-12 educational system as well as the universities. Strategic pressure points that can effect maximum change with the least investment of resources have to be identified. The statewide systemic reform programs have identified three major pressure points:

- Influencing policy changes that will catalyze reform.
• Nurturing and assisting schools and teachers in the adoption of rigorous standards and providing model curricula and effective teaching strategies designed to meet the standards.

• Professional development and empowerment of teachers and school administrators that will provide them with the skills and tools to incorporate the reform into their classrooms and to become agents of change in their schools or districts.

Some of the systemic initiatives have also identified and promoted the development of teacher preparation reform at the university level. In some cases, the initiatives have also developed an additional pressure point to reform Science and Mathematics courses in University Natural Sciences Departments. Although much harder to achieve, the reform of university level education is necessary to sustain the K-12 reform effort once the catalytic reform resources are gone. The university level reform is essential if a self-correcting education system is to be developed that is capable of making reform a continuous and adaptive process and capable of responding to an ever changing external environment.

A political strategy is needed in order to achieve systemwide reform given the limited reform resources and the need to produce institutional culture transformations. In order to act politically, it is absolutely necessary that alliances are negotiated, forged, and nurtured; this requires that reformers have political and negotiating skills in order to bring diverse groups together to achieve a common goal. Effective working alliances are diverse and idiosyncratic and vary from systemic initiative to systemic initiative; among the most valuable strategic alliances that have been forged are those between industry/business and policy makers; universities and school districts or departments of education; and schools with other schools or districts with districts. The alliances have helped to harness and optimize resources of the partners to achieve a common goal and influence reform.

Ultimately, systemic reform will work only if the existing educational system resources are harnessed and used optimally and effectively to implement the new educational paradigms that are tested and proven more effective and efficient in achieving student learning. All reform efforts start as experiments with new ways of doing things in an institution or in the classroom; they are then pilot tested and validated and, when proven to be more effective than previous approaches, they are widely disseminated. Only if both teachers and administrators buy into the new models, system resources will flow to implement the reform. Thus, the reform effort must have clear strategies for persuading and enticing or, if necessary, coopting the system to buy into the new paradigms. Only then system resources will be harnessed to scale-up the reform and achieve new and enduring effective ways of educating students. Harnessing system’s resources for reform is probably one of the biggest challenges of educational reform.
C. Systemic Reform is an Institutional Cultural Transformation Process

All social systems have well-defined rules and procedures by which they conduct their daily transactions or business which are based and guided by a set of beliefs and values which are both explicit and implicit. In other words, all social systems possess an institutional culture. But, when a system ceases to be optimally functional and reform is needed, it is precisely the institutional culture that can become the biggest impediment to change. Therefore, it is an imperative that all successful reform efforts address the issue of institutional culture and develop targeted strategies to modify the social transactions of the system needed to allow new and more effective ways of conducting the education process to be fully incorporated into the system and institutionalized.

Professional development of teachers and school administrators is one grassroots approach to achieve this goal. All of the systemic initiatives have recognized the importance of professional development and have put it to good use. Professional development takes many forms and traditionally includes one or more of the following elements: subject matter content, teaching/learning methodologies and strategies, new assessment methods, new organizational models, and new standards and model curricula. Yet, in its traditional form, professional development is essentially an individual process; the teaching/learning practices of a teacher are transformed through workshops, and then the teacher returns to an untransformed school, resulting in little or no systemic change. This approach fails to take into consideration that the objective of the reform is to produce an institutional cultural transformation which is a collective and systemic objective.

Several reform projects have directed their professional development to explicitly include the academic empowerment of teachers. Under these approaches, teachers master the new educational methods and strategies needed to take full control of their classroom teaching/learning environment and develop leadership skills to serve effectively as agents of change in their schools and/or districts. Another approach has been to bring teachers from entire schools or districts together and prepare them with skills and methods needed to form effective Communities of Learners. The implementation of the concept of community of learners in schools or districts is a powerful mechanism to achieve institution cultural transformation, to modify the behavior of these units, and to transform them into learning communities that can evolve and modify their beliefs and values to align them with the reform effort.

At the same time as the grassroots approach is implemented, a top-down approach is needed to promote the adoption of challenging Science and Mathematics standards, curricular frameworks, and enabling policies at the policy making levels of the educational system. This approach, coupled with the adoption of high-stakes tests based on the new standards and frameworks, constitutes a two-prong approach to
educational reform that will nurture and facilitate the necessary institutional cultural transformation. Another major challenge for the reformer is obtaining full alignment between these two approaches in order to achieve systemic change. The balance between top-down and grassroots approaches will vary from site to site depending on local political realities and opportunities; this balance may also change over time as the reform evolves.

D. The Need for a Unit of Change

Because of the sheer size and complexity of educational systems and the limited resources to achieve reform, the whole system cannot be initially tackled; hence the need for a unit of change. Many different units of change can, and have been, adopted by the systemic reform efforts. However, the concept of the unit of change has no practical value if the selected unit is smaller that the smallest unit in the system that possesses systemic features within itself. For this reason, many systemic programs have adopted the school or district as their unit of change.

The unit of change is the smallest unit in which systemic reform strategies are pilot tested and evaluated before they can be scaled-up in a meaningful manner to the rest of the system. This is the unit where the teaching/learning and institutional cultural transformations needed for successful educational reforms are pioneered. Therefore, the reformer needs to design a set of specific strategies tailor-made for this unit. Since the reform process, in order to be truly systemic and comprehensive, must address all the levels of the educational continuum, units of change are needed at all levels from K-12 to universities.

E. Scaling-Up

A strategy is needed to scale-up the reform horizontally to the other equivalent units of change within the system, once the unit of change has been defined and strategies developed, tested, and validated for their effectiveness in producing the desired reform goals and the necessary institution cultural transformation. The horizontal dissemination of the reform requires that multiple foci of reform be identified and developed. Various types and names have been given to these foci of reform, but we will use here the generic term Dissemination Centers to identify them. The Dissemination Centers serve the function of providing technical and academic assistance, building collaborations and, in some cases, they also provide professional development. All these services are provided to other units of change that buy-in and wish to join the reform process either in a voluntary manner or by specific directives from the educational system’s management. Motivation for units to join in the scaling-up process may come from new state policies and accountability systems that have been affected by the systemic reform. In this dissemination process, harnessing system’s resources for the reform becomes an increasingly important function of these Centers as new schools or districts join the reform under their leadership.
The biggest danger of the scaling-up process is that, as the reform reaches more units of change, the quality of the dissemination process tends to deteriorate and lose the quality and effectiveness that it had in the pilot stage. For this reason, assessment and quality control mechanisms must be put in place to monitor the scaling-up process and to ensure the ultimate success of the reform.

F. Using Assessment, Attribution and Accountability (A³) to Drive the Reform

Social systems are highly sensitive to external measurement and can react by modifying their behavior. This is particularly true of educational systems, which, due to political and public opinion pressures, are highly sensitive to student performance outcomes. For this reason, assessment can be made into a powerful tool to drive the reform process if it can be developed into a positive feedback mechanism.

The reformer must pay careful attention to three levels of student performance assessment: classroom assessment; assessment at the unit of change level; and systemwide assessment. The first is handled by providing new classroom assessment tools that are aligned with the standards and reform principles through professional development and then encouraging teachers to use them in their classrooms. The second is achieved by using existing system assessment tools or by developing special assessments to measure the progress of the unit of change in achieving reform goals. The third requires partnering, developing strategic alliances or influencing systemwide testing or assessment systems to align them with the new standards and reform goals.

In addition to student performance assessment, assessment tools need to be developed to measure teaching/learning cultural transformation in the units of change as well as other reform goals that are systemic in nature, such as organizational changes and the development of communities of learners.

Attributing improved student performance and the achievement of reform goals to reform efforts is not simple when dealing with complex systems that are buffeted by many initiatives and external forces. Because of the principle of stasis of social systems and the need to reform by persuasion and not by imposition, it is imperative that the reformer provides extensive evidence that the reform model produces an improved school environment and better results in student performance; thus, the issue of attribution is fundamental to achieve systemwide transformation.

The implementation of standards-based curricula and increased opportunities to learn are the elements of systemic reform that can be shown to have a direct impact on student achievement outcomes; this direct impact can be shown quantitatively by analyzing data and establishing correlations between these elements of reform. On the other hand, the impact of policy and convergence of resources on student outcomes is
indirect and much harder to establish because its impact is mediated by the implementation strategy chosen by the reform. In general, the effect of policy and convergence of resources on the implementation strategy can only be established qualitatively (see Figures 1a and 1b).

In order to make a convincing argument for attribution, a well-defined set of systemic metrics and student performance assessments must be designed, used, and converted to useful information. By triangulating all these results and developing a persuasive story line, the information and measurements can be used to attribute positive results to the reform and, thus gain allies and accelerate the reform process.

Accountability in an educational system is basically a prerogative of the system’s hierarchy, which is the one that controls and distributes system’s resources. The function of the reformer is to achieve changes in the accountability process to ensure that they are aligned with reform goals and produce positive and corrective feedback to drive the reform forward through persuasion and through the formation of strategic alliances with the industrial/business sectors and by influencing public opinion. Only when the system’s accountability process responds to assessment and recognizes the implications of attribution is the cycle of reform completed and institutionalization of the reform attainable.

G. Reform is a Process not a Goal: The Need for Self-correcting Systems

Changes in the external realities and forces create continuous pressures for reform. For this reason, reform is not a goal frozen in time but a continuous process by which functional systems adapt and adjust to respond constructively to external realities. Four elements, when properly aligned, can provide the appropriate feedback mechanisms to make the system more self-correcting: aligning systemic assessment with educational standards and frameworks; using the results of systemic student performance to guide professional development so that it responds to identified educational needs; developing and nurturing communities of learners at the school or district level so that institutional learning can take place; and using assessment and attribution to guide accountability and allocation of resources. At the grander scale, universities must come in to develop and sustain teacher preparation programs that will graduate teachers with the skills and knowledge needed to see reform as a process and capable of being life long learners. Only then is the full cycle of reform process completed.
Figure 1a

Correlation between Student Achievement and Systematic Implementation Attribution Model

- Hierarchical levels of influence
- Qualitative correlations between student achievement and implementation
- Changes and implementation between broad systemic changes and implementation
The Essential Difference Between Science and Mathematics (S&M) Reform Projects and Systemic Reform Process

Most Science and Mathematics (S&M) Reform Projects fail to have a fundamental and lasting influence on the educational system in which they are carried out. The ultimate reason for this is that the vast educational system’s resources are not harnessed to assimilate into the system the most effective educational curricula, teaching methods, tools or other products resulting from the reform project. As a result, the reform products and methods are not institutionalized. Great S&M projects come and go without affecting in fundamental ways the educational system’s culture; stasis prevails.

Many S&M reform projects fail to affect the educational system’s behavior because they do not provide effective feedback mechanisms to influence and change the system. The average S&M reform project usually has three major processal steps (see Figure 1): i) an analysis of deficient product or process in the system is identified and a proposal is written with a well conceived plan to correct the problem; then a successful proposal is sent to a foundation or agency for funding; ii) research and development of the project is done and then the pilot implementation of the reform strategies or products are tested with a control group; iii) the assessment of the impact of the reform products and strategies is done and if the results are positive, materials and products are prepared for dissemination and adoption by the educational system. The implicit and rarely explicit assumption is that if the products and strategies resulting from the project improve the system or solve a recognized problem in the system, the system will adapt, adopt, and institutionalize the improvements. As the S&T and business communities are well aware this is a rather simplistic view of the innovation process. Thus, it is no surprising that in most cases excellent project come and go and the system does not change!

As Figure 1 shows there are several missing steps and feedback mechanisms that must be in place if dissemination and adoption is to take place and if the behavior of the system is to change and make the necessary transformations needed to assimilate the new products and strategies so that innovation can actually occur. First, a systematic attribution analysis and design must be put in place with feedbacks to the assessment and the pilot phase in order to make a persuasive case that the new products and strategies are directly or indirectly responsible for the resulting improved teaching/learning environment. Attribution must be sufficiently persuasive to obtain buy-in from the upper policy and management element of the system and/or to persuade the units of change of the system, which for educational systems have to be the schools or districts. Only then sufficient resources from the system can be harnessed to proceed with a scaling-up phase.
As NSF discovered during the implementation of the Statewide Systemic Initiative (SSI) programs, attribution and scaling-up are not trivial processes and require considerable planning and sophisticated implementation strategies. Without scaling-up within the system units (schools and/or districts), the reform cannot become systemic in the sense of transforming the whole system. The scaling-up process can be seen as the entry point into the system’s organizational and operational structures and the first step in harnessing system resources to produce long-lasting innovation. From an S&T business innovation model point of view, these stages correspond to manufacturing, marketing, and selling of the innovation.

To complete the reform cycle, the system itself needs to bring into play several elements. The evaluation and institutional research element of the educational system that is being reformed must be robust enough to measure and detect the impact of the reform on the units that are participating in the scaling-up process. The accountability and allocation of resources process must be informed by the evaluation in order to promote change in other units of the system. Finally, the system-wide evaluation process including student performance tests and assessment of whole school performance, should be able to point towards new challenges and problems that can be tackled by new reform projects, thus completing the reform cycle (see Figure 1). If this complete innovation cycle is attained, the reform cycle is closed, and a self-correcting process will be in place allowing the system to correct itself. The system must understand that reform is a continuous never-ending process and not simply a correction goal.
The Self-Correcting Process: The Dynamo That Drives Successful Systemic Educational Reform

Figure 1

The System

- Assessment
- Attribution
- Scaling-Up

- Evaluation & Institutional Research
- Accountability and Allocation of Resources
- Evaluation
- Pilot Implementation of Reform Strategies

Strategic Plan For Reform

A

Pilot Implementation Of Reform Strategies
Establishing a Quasi-causal and Persuasive Relationship Between the Systemic Reform Initiatives and the Improvement of Student Learning: The Use of Assessment and Attribution to Drive the Educational Reform

Social systems are highly sensitive to external measurement and can react by modifying their behavior. This is particularly true of educational systems, which, due to political and public opinion pressures, are highly sensitive to student performance outcomes. For this reason, assessment can be made into a powerful tool to drive the reform process, if it can be developed into a positive feedback mechanism. In Puerto Rico, a self-correcting loop was developed to allow the system to correct itself through feedback by bringing together Assessment, Attribution, and Accountability to interact with each other: a process we called A³.

Assessment - Conventional approaches to evaluation are not readily adaptable to the evaluation of systemic educational reforms because of their uniqueness and complexity. A reform of a complex social system, to be effective, needs to tackle simultaneously multiple elements within the system. Building from the literature that suggests the use of multiple measures to triangulate results to identify trends, in Puerto Rico we used a combination of qualitative and quantitative approaches to measure the value-added of the reform initiatives.

Careful attention was given to three levels of student performance assessment: classroom assessment; school assessment, and systemwide assessment. In addition to student performance assessment, assessment tools were developed to measure the transformation of the teaching and learning culture at the unit of change (the school), as well as other reform goals that are systemic in nature, such as organizational changes, and the development of communities of learners.

Attribution – Attributing improved student performance and the achievement of reform goals to reform efforts is not simple when dealing with complex systems that are buffeted by many initiatives and external forces. In a system, the higher the synergy, the harder it is to disaggregate the impact of specific systemic strategies and activities, such as the S&M reform, from the overall systemic transformation. Nonetheless, we demonstrated that it is still possible to design metrics that can provide strong evidence of attribution to make a persuasive case for the decisive contribution of the S&M reform to the overall reform.

Accountability – Accountability in a social system is basically a prerogative of the system’s hierarchy, which is the one that controls and distributes system’s resources. Ultimately, systemic reform will work only if the existing educational system resources are harnessed and used optimally and effectively to implement the new educational strategies that were tested and proven more effectively and
efficient (attribution) in achieving student learning. Only when the system’s accountability process responds to assessment and recognizes the implications of attribution is the cycle of reform completed and institutionalization of the reform attainable.

**Measuring the Value-Added of the Systemic Reform in Puerto Rico: Improved Student Achievement**

To measure the impact of the reform on student performance pre and post-tests were developed for the fourth, eighth, and eleventh grades, using the public-released items from TIMSS, NAEP, and other sources. Two broad categories of items, both in S&M, were included in these tests: multiple choice items that test for mastery of content with a heavier emphasis on memory, and open-ended questions that were selected to measure higher order thinking skills within each discipline. Score distribution curves and scatter plots were used to look at individual student performance and overall performance of a school by aggregating the performance of all students of a given school.

Since the overall score distribution graphs for both S&M as well as for the different grade levels behaved similarly, we will only show a comparison of the *individual test score distributions* for Fourth Grade Mathematics (Figures 2 and 3).

Figures 1 and 2 illustrate that, on the average, there were more significant gains by all students thus validating the tenet that the reform is for ALL students. This tendency in gains was seen for both disciplines and across grade levels. But, even more significant is that the pattern of results where the distributions of the pre/post tests, for both multiple choice an open-ended questions, are skewed towards the higher scores of the scale, has remained consistent over three years. This skewness evidences the impact of the reform’s curricular and teaching/learning process in allowing students to express their full potential.

**Figure 1**

*Fourth Grade Multiple Choice:*

<table>
<thead>
<tr>
<th>Score Range</th>
<th>Number of Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 49</td>
<td>100</td>
</tr>
<tr>
<td>50 - 99</td>
<td>200</td>
</tr>
<tr>
<td>100 - 149</td>
<td>300</td>
</tr>
<tr>
<td>150 - 199</td>
<td>400</td>
</tr>
<tr>
<td>200 - 249</td>
<td>500</td>
</tr>
<tr>
<td>250 - 299</td>
<td>600</td>
</tr>
<tr>
<td>300 - 349</td>
<td>700</td>
</tr>
<tr>
<td>350 - 399</td>
<td>800</td>
</tr>
<tr>
<td>400 - 449</td>
<td>900</td>
</tr>
<tr>
<td>450 - 499</td>
<td>1000</td>
</tr>
<tr>
<td>500 - 549</td>
<td>1100</td>
</tr>
<tr>
<td>550 - 599</td>
<td>1200</td>
</tr>
<tr>
<td>600 - 649</td>
<td>1300</td>
</tr>
<tr>
<td>650 - 699</td>
<td>1400</td>
</tr>
<tr>
<td>700 - 749</td>
<td>1500</td>
</tr>
<tr>
<td>750 - 799</td>
<td>1600</td>
</tr>
<tr>
<td>800 - 849</td>
<td>1700</td>
</tr>
<tr>
<td>850 - 899</td>
<td>1800</td>
</tr>
<tr>
<td>900 - 949</td>
<td>1900</td>
</tr>
<tr>
<td>950 - 999</td>
<td>2000</td>
</tr>
</tbody>
</table>

**Figure 2**

*Fourth Grade Open Ended Questions:*
Figures 3 and 4 illustrate the **average score distribution by school** for multiple-choice and open-ended questions in Mathematics for the Fourth Grade. When Figures 1 and 2 are compared with Figures 3 and 4, it is evident that the shift in the mean scores for schools is significantly greater than that for individual students, showing differences of at least one standard deviation when compared with the mean of the pre-tests. These results make a strong case for the fact that the reform is achieving its intended goal, because, on the average, participating schools have institutionalized a teaching and learning culture that promotes a positive learning environment for ALL students.
Figures 3 and 4 also show that the mean score gains between the pre and post tests were higher for the open-ended questions (depth of understanding and thinking skills) than for the multiple-choice questions (content and memory). Moreover, the distribution of the scores in the open-ended questions is even more skewed towards the high scores than the distribution of the multiple choice questions since some schools reached a mean of 800 and higher in the post-test.

Figures 5 and 6 show the scatter plot between mean school scores in the pre/post tests for the multiple choice and open-ended questions for Fourth Grade Mathematics. These figure shows that the majority of the schools made significant gains in the post tests since the points of the plot are located above the diagonal with few exceptions. According to Dr. Edward Kifer, the only country that has shown patterns of student achievement where student performance is higher in depth of understanding vs. content (i.e. memory) items was Japan.
Figure 5
Fourth Grade Multiple Choice:
Math Means Pre-Test and Post-Test by School
2000-01

Figure 6
Fourth Grade Open-Ended:
Math Means Pre-Test and Post-Test by School
2000-01
The results shown in the scatter plots are compelling evidence of the accomplishments of the S&M reform in developing and adopting standards-based curricular and student-centered constructivist teaching and learning methods that promote depth of understanding within each discipline. Scatter plots for all other grade levels and for both S&M show similar general patterns of performance.

Another metric used to measure the value-added of the reform were the scores obtained in the College Entrance Examination Board (CEEB) tests, the equivalent of the SAT test in the United States. When the average scores of the schools (schools that have been implementing the reform during at least three years –i.e. full implementation; schools that were still experimenting with the strategies and the curricular content promoted by the reform –i.e. partial implementation; private schools, and non-participating public schools) were obtained, results showed that the gap in performance between students from schools that had been engaged in full implementation and students from private schools (middle and upper class students) was essentially eliminated (See Table 1). The table shows that the average scores of the schools engaged in full implementation exceeded those of the private schools by 5 points in the Mathematics Reasoning Test and by 10 points in the Mathematics Achievement Test.

<table>
<thead>
<tr>
<th>Type of High School</th>
<th>Mathematics Reasoning</th>
<th>N (Students)</th>
<th>Mathematics Achievement</th>
<th>N (students)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Implementation</td>
<td>547</td>
<td>155</td>
<td>553</td>
<td>155</td>
</tr>
<tr>
<td>Partial Implementation</td>
<td>461</td>
<td>155</td>
<td>472</td>
<td>155</td>
</tr>
<tr>
<td>Private Schools</td>
<td>542</td>
<td>6036</td>
<td>543</td>
<td>5379</td>
</tr>
<tr>
<td>Non-Participating Public Schools</td>
<td>457</td>
<td>21149</td>
<td>458</td>
<td>20536</td>
</tr>
</tbody>
</table>

Table 1 also shows that students from schools engaged in full implementation outperformed students from non-participating schools by 90 and 95 points respectively. In contrast, students from schools engaged in partial implementation outperformed students from non-participating public schools by 4 points and 14 points respectively. These findings emphasize the need for a full implementation of the reform before significant changes in student performance can be obtained. Also, the significance of these results is that students from low-income families can now successfully compete for admission to Puerto Rico’s most competitive universities.
Above all, these results provide the most compelling evidence of the clear impact of the S&M reform on the overall educational system, since the longer the reform has been in place in the school, the more dramatic is its impact on student performance of the school relative to schools that have not yet entered the reform process. It would be difficult if not impossible to explain what has happened in these schools if not for the involvement in the S&M reform.

As has been shown, the closed positive feedback loop between assessment, attribution and accountability was a powerful tool to drive the reform process, and was essential to achieve a long lasting reform as the system assimilated the innovative products and strategies to improve the teaching/learning environment. The demonstrated positive results obtained in the reformed schools (assessment and attribution), was instrumental in getting other schools to enter the reform during the scaling-up phase. The incorporation of a responsive accountability system, on the other hand, ensured that the rest of the schools in the system eventually join the reform. The schools that went through the reform process following the whole-school approach did sustain the teaching/learning innovations after the reform project ended, evidencing the long-lasting impact of the reform.
Validating the Transferability of a Systemic Reform Model

The first and most important challenge that any reformer faces in implementing a successful and lasting systemic Science and Mathematics reform, is to understand how complex social systems operate, the nature of the system’s interactions and feedback, and how to influence its behavior. After experimenting with intervention programs for more than ten years, we accepted this challenge and designed a systemic reform model based on the theory of social change in complex social systems. This, in turn, kept the reform clear of becoming culturally specific, and opening the possibility for the model to be transferred and adapted by other educational systems.

The Puerto Rico’s Science and Mathematics reform was successfully implemented in 750 public schools, more than half of the public schools on the Island, where 80% of the students come from low-income families and show many other characteristics of inner city schools in the United States. The S&M reform implemented significantly improved the academic performance of underserved students vis-à-vis the performance of the more affluent students from Puerto Rico’s private schools (see paper on Measuring the Value-Added of the Systemic Reform in Puerto Rico). Private schools in Puerto Rico are more akin to suburban schools in the mainland in the socioeconomic profile of their students and in their access to resources.

Thus, the next step was to test the hypothesis if this systemic reform model for schools with underserved populations was transferable to inner city schools in the mainland and if it would be as successful in improving student performance in Science and Mathematics.

The New York City Experiment

The Puerto Rico’s systemic reform model was tested in 12 New York City Public Schools in the South Bronx (schools with large numbers of Puerto Rican students) through a three-year project (1998-2001) entitled “Puerto Rico/New York City Educational Linkages Demonstration Project”. Using the school as the unit of change, an intensive standards-based professional development program tailored to the needs of the teachers and schools, was designed and implemented following the “Whole School Approach”. The other key elements of the reform: 1) standards based-curricula; 2) low-cost materials; 3) use of authentic assessment strategies; 4) alternative management strategies, and 5) supportive evaluation and monitoring, were adapted, taking into consideration the nature of the New York’s system. The central node of this project was the creation of three Pilot Development Centers in intermediate schools modeled after the Puerto Rico Regional Professional
Development Centers to transform the teaching and learning culture of participating schools.

Assessment instruments to measure student learning over the course of one year were designed similar to those used in Puerto Rico. Tests were designed for Science and Mathematics using public-released items from TIMSS and NAEP among other sources. Results in these tests were comparable to those of students from Puerto Rico who tended to perform better in open-ended questions than in multiple-choice questions. Patterns of gain were similar to those of 8th grade students who took the test in the 2000-01 school year.

### New York Schools: Seventh Grade Average Student Scores in Multiple Choice Pre/Post Tests 2000-01

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Administration</th>
<th>N</th>
<th>Minimum Score</th>
<th>Average</th>
<th>Maximum Score</th>
<th>Standard Deviation</th>
<th>Net Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science</td>
<td>Pre-Test</td>
<td>331</td>
<td>136</td>
<td>435</td>
<td>617</td>
<td>78</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Post-Test</td>
<td>297</td>
<td>121</td>
<td>455</td>
<td>648</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td>Pre-Test</td>
<td>322</td>
<td>246</td>
<td>430</td>
<td>626</td>
<td>68</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Post-Test</td>
<td>291</td>
<td>224</td>
<td>468</td>
<td>741</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>

### New York Schools: Seventh Grade Average Student Scores in Open-Ended Questions Pre/Post Tests 2000-01

<table>
<thead>
<tr>
<th>Content Area</th>
<th>Administration</th>
<th>N</th>
<th>Minimum Score</th>
<th>Average</th>
<th>Maximum Score</th>
<th>Standard Deviation</th>
<th>Net Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science</td>
<td>Pre-Test</td>
<td>331</td>
<td>168</td>
<td>440</td>
<td>714</td>
<td>89</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Post-Test</td>
<td>297</td>
<td>142</td>
<td>477</td>
<td>766</td>
<td>107</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td>Pre-Test</td>
<td>322</td>
<td>258</td>
<td>425</td>
<td>692</td>
<td>84</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Post-Test</td>
<td>291</td>
<td>240</td>
<td>470</td>
<td>764</td>
<td>105</td>
<td></td>
</tr>
</tbody>
</table>

**Lessons Learned:**

1. The complexities of the largest urban school system in the United States were evident throughout the implementation of the project. These had to be taken into consideration in order to implement the systemic reform model.
a. A strong local base of operations is required so that appropriate follow-up can be given to teachers not only in the classrooms but with their educational leaders.

b. The level of teacher mobility within and across districts becomes a challenge in terms of pertinent professional development and follow-up as well as measurement of student and teacher gains. High student mobility posed a similar challenge.

c. Teachers who are teaching Science and Mathematics to bilingual students, in general, need to strengthen their mastery of content areas since most of them have degrees in other fields.

d. Local colleges and universities need to become more engaged in K-12 education, so they can support teachers in their daily teaching and learning practices.

2. When teachers found a strategy that helped them meet the needs of their students, such as the standards-based content and methodology, they adopted the strategy as their own. For instance, some teachers were assigned monolingual students after they had been participating in the project, and they transferred the reform strategies to these groups.

In conclusion, the pilot test carried out in the New York schools showed that the reform model can be successfully transferred to a different social and cultural environment, if the key elements of systemic reform are implemented taking into consideration the specific complexities of the new system.