SUSTAINING CHANGE: A STUDY OF NINE SCHOOL DISTRICTS WITH ENDURING PROGRAMS

DR. JEANNE ROSE CENTURY AND MS. ABIGAIL JURIST LEVY

CENTER FOR SCIENCE EDUCATION
EDUCATION DEVELOPMENT CENTER, INC.
55 CHAPEL STREET
NEWTON, MA 02458
617-969-7100 x2414; x2437

Paper Presented at the Annual Meeting of the
American Educational Research Association
April 3, 2002

The authors would like to acknowledge the contributions of staff members of the Center for Science Education (CSE), Education Development Center, Inc., Newton, Mass., and the Caltech Pre-College Science Initiative (CAPSI) in Pasadena, Calif.

This research was supported by a grant from the National Science Foundation. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

STUDY PURPOSE AND OVERVIEW

The question of how to sustain change in education is an enduring and unanswered one. Theories about change and its viability are widespread, but few researchers have conducted field-based research on the question of sustainability of reform. Therefore, there are few concrete, illustrative findings on which researchers, educators, and theorists can base valid interpretations, actions, and theories. Even less research has focused on districts working to systematically improve their science education programs. This study begins to address the absence of such studies by engaging in field-based research to identify and document factors in school systems that contribute to sustained educational change in science education.

The *Researching the Sustainability of Reform* (RSR) project focused on the question of how to maintain the gains of an initial educational change process and support continuing reform over time. Within the broader study of sustainability, the research paid particular attention to systemwide approaches to science education reform as well as the role that external funds can play in initiating reforms that are sustained. The research was conducted by staff of the Center for Science Education at Education Development Center, Inc. (EDC), in Newton, Mass., in collaboration with staff at the Caltech Pre-College Science Initiative (CAPSI) in Pasadena, Calif., and was supported by a grant from the National Science Foundation.

The primary purpose of this study was to provide districts now engaged in improving their science education programs and districts that are considering doing so in the future, with information that can help them more strategically and effectively build an infrastructure for long-term improvement. Specifically, this study focused on nine communities with K–6 science education programs begun from nearly 10 to 30 years ago. These communities differed in their sources of funding as well as the longevity of their programs. This study investigated how, and the extent to which, these communities have sustained their science education programs and the factors that have contributed to this sustainability.

BACKGROUND AND THEORETICAL FRAMEWORK

Over the history of our developing education system we have seen the rise and demise of many reforms in science education. In the first half of the 20th century, for example, science education change efforts co-evolved with developing social theories and ideas about cognitive development and learning theory. The "Project Method" approach in the early 1900s attempted to make learning more experiential and meaningful for students. The 1920s and 30s brought more "progressive" child-centered approaches (Portelli, 1987). Social meliorism soon followed, using the curriculum to bring about improved social conditions and later, the 1940s' "life adjustment" movement focused on "preparing students for life" (DeBoer 1991). Though some of these ideas became popular, they were never strategically implemented as comprehensive education reform plans and, thus, made no widespread lasting impact on science instruction. In most classrooms, between 1920 and 1957, children continued reading about science instead of doing science, and the "curricula became somewhat fossilized" (McCormack, 1992).

In the second half of the 20th century, science reform efforts grew out of several major issues of national importance. During World War II, concerns grew about inadequate general science education, drainage of the pool of scientists in the field, inadequate teacher education, and the importance of science to our national security. The National Science Foundation (NSF) was established, laying the groundwork for the NSF-funded curriculum efforts of the 1950s and 60s that followed the launching of Sputnik in 1957. These programs "focused on reflecting the nature of science, as seen by practicing scientists, and on learning by inquiry and discovery" (McCormack, 1992). Some of these programs were used at one time in up to 25 percent of the districts in the country; and one, BSCS, was used in almost 50 percent of school districts

(DeBoer, 1991). Significantly, however, few places *continued* using the programs beyond the first few years following introduction.

As the 1960s progressed, the field of science education began to move toward stressing environmental education. Then, according to DeBoer, the watchword of the 1970s became "scientific literacy," meaning a focus on relevance to the lives of students and social issues (1991). And the 1980s brought a shift toward the role of science education within the broader theme of science-technology and society (McCormack, 1992; DeBoer, 1991). Until the late 1980s, the reform efforts in science education continued to focus almost exclusively on curriculum development and, to a lesser extent, on teacher workshops. Then, in the late 1980s and 90s, reform efforts turned to setting standards and "systemic change."

This relatively brief history of science education in this country illustrates the ebb and flow of reform ideas and their implementation life spans. Given the short time that the field of science education has existed in this country, we would be hard pressed to develop a solid argument that *any* science education reform effort had adequate opportunity to become fully realized. Perhaps this is why, after more than 100 years of reforms—*all* of which turned away from traditional approaches—much, if not most, science still is taught using traditional texts with traditional methods.

Concurrent with the evolution of science education over the last quarter century, a body of literature developed throughout the 1970s and 80s that established some basic understandings about why change efforts have not had long lasting effects (Loucks & Hall, 1977). One landmark study took place in 1973–1978, when the RAND corporation funded an investigation of four federally funded reform programs of the 1960s. Although each of these programs had widely varied purposes and target audiences, they were similar in that they were intended to bring about innovations in educational practice by using temporary seed moneys to initiate and implement the programs (McLaughlin, 1989). The findings suggested that although the programs achieved isolated success, they shared some common problems that prevented their efforts from reaching wider audiences.

Two other studies of that time demonstrated that even when projects succeeded on a small scale, they failed to have wide-scale impact. The Pilot State Dissemination Project (1972), for example, demonstrated successes, but only with prohibitively expensive ongoing technical assistance. In contrast, a study on Project Innovation Packages (1975) demonstrated that an approach based on the absence of support can be just as problematic. In this case, materials were specifically designed to be "self-teaching" but had little lasting effect on classroom practice because they lacked the requisite supports from other parts of the system (such as professional development) (Sashkin & Egermeier, 1992).

These studies demonstrated that individually designed and implemented programs focusing on single issues or single inputs could indeed exhibit small successes, but that these benefits existed in isolation and in highly variable locations. They often ignored "the systemic and interconnected conditions that influence classroom practice" (McLaughlin, 1989). John Covaleskie argues that the nature of systems is such that they sustain mediocrity and strive for efficiency. Therefore, as reforms "diffuse through the system they tend to become less reforms as they are modified to conform to the systemic demands for efficiency" (1994). Sarason contends that the "regularities" in school systems powerfully endure and deter the sustainability of any curriculum or structural changes. He argues that if these regularities are not targeted, change will never be sustained. If they are targeted, Sarason suggests, a domino effect of changes can result (Sarason, 1996).

Now practitioners are asking how to hit these "targets" so they can create an organizational structure that facilitates ongoing adaptation and growth and supports and sustains progress already made. If they continue to operate as usual, their efforts are likely to fall along the path of previous failed reforms. Drawing from the literature, this study sought out the answers to their questions by looking at districts that had enduring hands-on science education programs, identifying ways the leaders of those programs focused their efforts and investments, and distinguishing the factors that influenced the outcomes of those actions.

METHODS OF INQUIRY AND DATA SOURCES

In 1998, the Center for Science Education at EDC and its research collaborator, the Caltech Pre-college Science Initiative (CAPSI), set out to better understand the sustainability of hands-on, elementary science programs. We undertook a study of nine districts that had had such programs in place for 7–30 years or more. The districts were diverse in size, location, demography, and resources (see Table I), but shared a commitment to the use of science instructional materials made available to teachers and their students in the form of kits. In brief, science kits focus on one or two core science concepts and contain enough materials and lesson plans for teachers to explore these concepts with their students over a six- to nine-week period. Typically, a science materials center delivers the fully stocked kits (usually one or two large plastic tubs) to classrooms and retrieves them at the close of the unit, when they refurbish and prepare them for delivery to the next set of classrooms. On average, a class uses three to four kits during the school year, and kits rotate from one set of classrooms to the next.

The study was guided by the following global research question: What contributes to or inhibits the sustainability of a districtwide, hands-on inquiry science program? We also included the following sub-questions in our investigation:

1) What is the current status of the science program and how does it compare with the initial goals and implementation of the program?
2) What conditions and contexts surrounding a science education reform effort impact its sustainability?
3) What decisions have practitioners made and what strategies have they used to bring about enduring change and build capacity for continuous growth?
4) How has the capacity of the practitioners in the system and the capacity of the system itself affected the sustainability of the reform?
5) What is the role of external funds as a catalyst and/or support for lasting, widespread reform?

To address these questions, the research design employed primarily qualitative methods, supplemented with a survey analyzed using quantitative approaches. Data collection focused on documentation of past events, understanding the current status of the science education program, and identifying the relationship between the science education program and the larger school system. The research team gathered a range of data from the following sources: state and district documents, teacher and principal surveys, interviews, and field notes from observations of classroom instruction and professional development. The data focused on the following areas: program goals and rationale, curriculum, professional growth, physical resources, leadership, financial resources and management, accountability and assessment, communication, partnerships, internal and external contexts, and district capacity.

The research was initiated in August 1998 and concluded in July 2001. During each year of the three-year project, research teams made week-long visits to the districts. During the site visits, researchers conducted semi-structured interviews with classroom teachers, principals, central office administrators, science program leaders and external program partners, community members such as school board members, and others. We conducted approximately 50–60 interviews in each district, in addition to interviewing program leaders several times during each visit. We made approximately 20 classroom observations of science instruction, and observed two to four sessions of professional development relevant to the elementary science program. Documents important to the program—instructional materials, science standards, program descriptions, goals, progress reports, and articles in the press—were reviewed. Finally, we conducted two informal surveys: one targeted all of the elementary principals in each district and the other targeted a random sample of 100 elementary classroom teachers in each district.

As we interpreted our data with an eye toward ensuring that our findings would be useful for practitioners, it was important for us to keep several caveats in mind. First, our intention was to explain how and why the programs we studied had endured, rather than predict the likelihood of their future sustainability. Furthermore, we did not set out to create a model of design or implementation to be used by other districts. Although we looked for common themes across sites that might reveal such a model, we found that the wide variations in contexts and conditions in each site necessitated that no single model could lead to success in multiple sites. Second, since we focused our attention only on programs that had been sustained (and did not look at failed efforts), we are unable to draw comparisons between the two groups. It would be incorrect for us to assume that the districts and programs we studied are significantly different from those where programs did not last. What we can do, however, is highlight the themes that were consistent across our sites and draw lessons from the characteristics, approaches, and outcomes that were common among them.

FINDINGS

Maintenance Versus Sustainability

Before looking closely at the factors that contribute to and inhibit sustainability, it is important to clarify what we mean by "sustainability." Early on, we recognized that educators commonly viewed program sustainability as program maintenance—embedding a program, as designed, into a standing operating system. As we explored the meaning of sustainability in the context of these nine school districts, we found that this view of sustainability was limiting. We acknowledged that we were coming to understand "sustaining districtwide education reform" as a contradiction in terms because at the same time that school districts want to maintain the innovations they put in place, they also need to continually adapt and improve them. The tension between maintenance and adaptation grew to be at the heart of our research as we sought the answers to two questions: (1) Was the program that we saw essentially the same one that had originally been implemented, a near or distant relative, or one that was virtually unrelated to the original? and (2) What factors had contributed to the program's endurance and adaptation(s)?

As the project progressed, we began to make clear distinctions between programs that had been maintained and those that had demonstrated sustainability. A program could be considered maintained if its base elements (e.g., instructional materials, professional development program, leadership plan) were well established and were commonly accepted as standard practice. We described sustainability, on the other hand, as the ability of a program to withstand shocks over time by maintaining core values and beliefs and using them to guide its adaptations to change. In other words, a program must be maintained before it can reach a stage of sustainability, but it cannot be stalled at maintenance; it must develop an ability to evolve and adapt.

> ***Sustainability: The ability of a program to withstand shocks over time while maintaining core beliefs and values and using them to guide its adaptations to change.***

The need for adaptability is critical for two reasons. First, changes occur regularly within districts and exert pressures on educational programs. These changes may offer opportunities or they may pose obstacles. Either way, program and district leaders must address them. Examples of such changes, or what some might refer to as "critical events," include a change in district or program leadership, a shift in political agendas within the school district or broader community, a budget crisis or change in district priorities, a large turnover in teaching staff, or a curriculum adoption.

Given these circumstances, it is challenging enough to simply initiate and maintain a kit-based program. Sustaining a program in the long term is even more complex not only because introducing ideas to, and establishing new practices with, teachers is difficult, but also because science presents unique challenges. It is not uncommon to find elementary teachers reluctant to teach science because they lack familiarity with the science content. Or, even when they are comfortable with the content, they may be overwhelmed by the need to manage the range of materials and activities in a science lesson. Additionally, in today's environment for accountability for reading and mathematics, elementary teachers feel pressure to give classroom attention to those subjects, reducing or even eliminating time for science instruction.

The challenge of supporting a districtwide science program under these circumstances is compounded by the fact that there is minimal, if any, accountability to ensure that teachers actually teach science; teachers can easily avoid teaching a science program altogether. Program leaders must, and do, work with teachers on essentially a voluntary basis, reaching out to them and their principals through means devised to attract them and appeal to their interest in providing a full and rich educational program. They must find compelling ways to reach teachers, principals, and administrators and lead them to embrace, as we stated in our definition of sustainability, the "core beliefs and values" of the program.

MOVING TOWARD SUSTAINABILITY OVER TIME

Before moving into descriptions of the factors we identified as contributing to or inhibiting sustainability, it is essential to understand one more point. Each sustained program must progress through what we identified as three stages of development: establishment, maturation, and evolution.

The *establishment* phase focuses on introducing the use of kits for the instruction of science, distributing the kits to the teachers, and implementing professional development programs for teachers and administrators. This stage requires that leaders focus on the very concrete elements of the program, making sure that they are well established, accepted, and working efficiently and predictably across the district. In addition to introductory professional development, leaders also must focus on the system for distributing science kits, ensuring that they will be collected, refurbished, and redistributed efficiently and dependably to all schools and classrooms. Leaders also must establish lines of communication and support systems so that teachers will be able to report their concerns and so leaders can understand and respond to problems as they arise.

The next developmental phase of a program is *maturation* and focuses on embedding the use of kits across the district and arriving at a point where kit use is habitual, even in the absence of the limelight that accompanies a "new" initiative. In this phase, the program itself is fairly secure in its structure although not immune to the events that may pose a threat to the program's existence. On the whole, the program has been established and accepted, the materials system functions fairly smoothly, and professional development for teachers is available (most often on a voluntary basis), at least for teachers new to the district. Program leaders must convince teachers, principals, and central office administrators that the science program still requires their attention to improve and advance, and they must compete for attention and devise ways to keep the importance of science in the forefront. Leaders must work to continually assess the nature of the program's imperfections and endeavor to address them.

The third phase is *evolution*. The hallmarks for the evolution phase are growth and improvement. Even though the programs are still not immune to threats and far from perfect in structure or instruction, leaders in this phase seem to construct for themselves a more intangible set of goals. They still must work hard to understand the current status of the program and address the recurring challenges of resources, materials management, and professional development. But they are also concerned with developing and refining teachers' understanding of science content and pedagogy. They are interested in advancing a level of appreciation and understanding among their teachers and within the district's administration that will enable their programs to absorb new ideas and research findings regarding elementary science instruction. During this phase, program leaders focus on helping teachers develop a deeper understanding of their capacity to teach both science concepts and the scientific process.

In each of these three phases in a program's development, the program has particular goals and particular strategies for pursuing those goals. Moreover, program goals have important implications at all levels of a school system, meaning at the classroom level, the school level, and the district level. To be sustained, program goals must be realized at different levels, which requires multiple strategies often employed simultaneously by program leaders. Thus, at any given point in the development of a program, program leaders might direct their attention to particular influential factors at any of these different levels of the system (see table below). Together, the phase of development and the program leader's level of orientation determine the importance and priority of each of the factors described below.

	Establishment	**Maturation**	**Evolution**
District			
School			
Classroom			

Each factor can appear in any cell depending on context, circumstances and priorities.

Influential Conditions and Factors

The findings of the study include nine site summaries and a cross-site report. Each of the nine site summaries is primarily descriptive and provides the reader with an account of the origins, implementation, and evolution of one of the nine science programs studied. Each summary also offers a brief analytic section that identifies how a subset of the factors described below were of particular relevance in that site. The cross-site report draws from all nine sites to identify common themes and recurring issues relevant to sustainability. It is primarily analytic while offering concrete supporting examples from the nine study sites. The cross-site report also includes suggested strategies for districts wishing to analyze and improve their science education reform efforts.

The remainder of this paper summarizes the main findings of the cross-site report. The findings are divided into three sections. The first entails the influential conditions—the context and environment—in which the factors that contribute to or inhibit sustainability operate. The second section includes those factors that are related to concrete elements of a science education program, such as leadership, professional development, and curriculum. The third section includes factors that are less tangible and relate to influences on a program and outcomes that result from the program. The conditions and factors are not mutually exclusive in their influences on sustained programs. There is a great deal of iterative overlap that one should not overlook even though the findings are organized in a linear manner. With this caveat, this paper defines each of the contextual influences and factors and highlights two of those factors in detail. For a more complete "map" of the findings, see the Appendix.

Influential Conditions

Community and district features weigh heavily on the ultimate success or failure of reform efforts that operate in those places. We identified three aspects of this context, that emerged as particularly noteworthy in supporting or inhibiting the development of the sustained programs. The summaries below highlight the aspects of these issues that most clearly relate to sustainability.

1) School System Culture

School system culture encompasses the nature of the human, structural, and systemic environment in which the science program has to function. Specifically, the human environment refers to the number and efficiency of communication channels between individuals in the system and the extent to which individuals are encouraged or supported in their efforts to work together in a collegial manner. The structural environment refers to the organizational hierarchy and how strict or formal that hierarchy actually is. Finally, the systemic environment refers to accepted and/or expected practices throughout the system. These include actions such as participating in professional development outside of school hours with a minimum of compensation (volunteerism), support for professional growth, and extent of support for innovation. Leaders in sustained programs do not underestimate the power of culture. It sets a foundation for the ways and extent to which the other factors named below contribute to and inhibit the sustained program. When there is the will to and interest in supporting a hands-on program, efforts to bring that program to fruition must be compatible with the culture or, even though well intentioned, they are likely to fail.

2) Decision Making and Power

Decision making and power refers to both the formal and informal means by which decisions are made in a district. Understanding this structure entails discerning where the power ultimately resides and who has access to it. Leaders of sustained programs understand this structure as well as how to negotiate their way through it to be able to influence choices that might affect the science program. In general, leaders of sustained programs have relatively little direct control over many potential influences on their programs. Decisions about when and how to change the science program are made at multiple levels with a single person at any one of those levels capable of impeding progress unless appropriate precautions are taken. Leaders of sustained programs then need to work to gain access to decision-making authority normally out of their reach through developing personal relationships and strategically building support.

3) Science for All

"Science for All" is a phrase that often refers to the need to narrow the access and opportunity gaps between differing constituencies, such as those defined by gender, SES, or race/ethnicity. In this study, "science for all" refers to that need, as well as to how, and to what extent, districts ensure access to the science program for all students. This issue can be a motivating force for systemwide reform and provide a strong rationale for those interested in garnering support for a districtwide science education program. Some of the sustained programs addressed it explicitly, but more often, it was embedded in larger district mission statements. The actual relationship between "science for all" and whether a district's hands-on science program is ultimately sustained is deeply rooted in the day-to-day work of the project in subtle, hidden, and mostly unspoken ways.

Factors that Are Related to the Concrete Elements of a Science Education Program

1) Accountability

Accountability measures, by their presence and absence, are a significant factor for sustained programs. We have defined accountability in two ways: (1) as a measure of student learning and (2) as a means of holding teachers and administrators responsible for the implementation of the science program. Mechanisms for accountability of the first type would include student written and performance tests, student work, and writing in student science notebooks. Mechanisms for accountability of the second type would include supervisory procedures, observation guidelines, and stated district learning goals. The sites in this study used (and omitted) a range of accountability measures, the presence (or absence) of which portray curricular, personal, and political priorities at various times in the evolution of the sustained science program.

2) Instructional Materials

Instructional materials are an essential component of any science education program. In the programs studied, all of the instructional materials were primarily kit-based, meaning they were comprised of bins or boxes that include a teacher's guide and the necessary manipulatives for teaching the lessons outlined in that guide. The issue of instructional materials encompasses a number of components that contribute to or inhibit sustainability. These include selection process, materials organization and refurbishment systems, and processes for adapting the materials. The decisions that leaders make in each of these areas determine the acceptance, use, and stability of this essential component of a hands-on elementary science program.

3) Leadership

In sustained programs, leaders' influences and supports are wide ranging and evident at all levels in the system. They extend from formally identified leaders (e.g., district science coordinators, fully released science resource teachers, and school site liaisons) to informal or "behind the scenes" leaders (e.g., school board members, assistant superintendents for curriculum, and community members). Successful leaders of all types are flexible with the varying needs of the program over time and respond to shifting district conditions. Furthermore, leadership styles and strategic choices in sustained programs are highly responsive to district culture. For example, science coordinators' approaches may vary from those that are aggressive and "front-line" to those that operate with a more subtle "behind the scenes" tack. Some focus on centralized approaches to reform, while others cultivate a more grass-roots style. In any case, strategies that are successful echo characteristics of the district's human, organizational, and systemic culture.

4) Money

Many equate sustainability with district financial commitment for a program. There is no question that money is a critical player in a sustained program, but its role is far more complex than the simple presence or absence of financial resources. The source of the money, the way it is used at different points in the evolving life of the project, and the nature of standard district practices and interactions with regard to money all are significant issues. Some sustained programs had large external grants during the establishment and maturation phases. While benefiting from the wealth of increased resources at that time, they faced new and sometimes unforeseen challenges, such as internal competition, lack of consideration for other resources, and constraints on their reform program design. Other programs with little inclination to seek outside support enjoyed the benefit of deep internal commitment while also facing difficulties, such as isolation, internal vulnerability, and limitations in the potential scope of work. Over the long term, sites with large influxes of financial resources and those that developed incrementally on their own demonstrated that both paths offer opportunities and obstacles to the development of a sustained program. Depending on where that program is in its evolution and the contexts and conditions influencing the program at that time, the amount and source of funding can both help and hinder sustainability.

5) Partnerships

Partnerships, while contributing to sustained programs in many ways, also can carry a heavy cost or even be a burden to those programs. In this study, partnerships fell into two broad categories. There were what one might call "deep" or comprehensive partnerships that required investments of resources and political currency and entailed shared planning and leadership. Such partnerships occurred mostly at the district level and were rare. More common were more "limited" partnerships at the school level that, while somewhat superficial and supplemental, still enriched the program. These partnerships most often took the form of sponsorships for resources, such as science materials or assistance with carrying out school meetings.

6) Professional Development

Professional development in the context of hands-on elementary science programs refers to activities focused on increasing teachers', principals' and administrators' capacity to understand and implement hands-on, inquiry-based science in their classrooms or schools, grasp the scientific content of particular units or lessons, and manage materials and students interactions with those materials. Such activities might include trainings on kit use (mandatory or voluntary), summer academies focusing on inquiry and/or content, study groups entailing individual exploration of science questions, and follow-up debriefings on how the kits worked in the classroom. In the absence of clear data on the impact of specific professional development activities on classroom practice and/or student outcomes, we explored several other avenues for understanding the role of professional development in sustained hands-on elementary science education programs.

One important issue is the relationship between the professional development's point of introduction (where the program is in its development) and the duration of its impact in the sustained program. For example, while it is clear that kit training sessions are essential in the establishment phase of the program, during the later maturation and evolutionary phases, such sessions haven't been consistently widespread nor have they necessarily been well attended. Still, the programs continue to endure, although sometimes haltingly. Another important issue is the level of perceived need for professional development among teachers, administrators, and program leaders. When there is a lack of congruence between offerings and perceived need resources are wasted, yet the programs still can last. And a third critical lens for understanding professional development is the role it plays in introducing, illustrating, indoctrinating, and disseminating the program philosophy.

<u>Factors that Relate to Influences on a Program and Outcomes that Result from the Program</u>

1) Critical Mass

Typically, discussions of critical mass focus on numbers: numbers of teachers participating, numbers of students reached, and the resource staff to teacher ratio. This is consistent with a view that a prerequisite for a sustained hands-on science program is that a minimum number of teachers teach hands-on science, thus making it, in practice, the standard for the district. Given our definition of sustainability, we have expanded on this description of critical mass to say that the program needs enough participating teachers to create a *culture of program self-generation*. Furthermore, we have found that in sustained programs, "critical mass" can encompass a range of ideas that develop beyond identifying a magic number of teachers. One alternative approach, for example, is to think about critical mass through the lens of perception (more on perception follows). More specifically, the truly important factor contributing to sustainability of a hands-on program may not be the actual number of teachers teaching the program but, rather, the number of teachers who are *perceived* to be teaching the program. Another alternative perspective on critical mass comes from the fact that it usually refers to the breadth of a program—numbers of people using the program. In sustained programs, it also is essential to achieve critical mass of the depth of the program—numbers of people *understanding and committing to* the philosophy and goals of the program.

2) History

History, as it relates to sustained programs, is comprised of two components: origins and longevity. The origin of a program is composed of the *circumstances* that were in place and the *events* that occurred from the initial conceptualization of the science program to the beginning of the establishment phase. Examples of such circumstances include the qualities and background of the early leader(s), the motivation and/or catalyst for beginning the program, extent of support from central office leadership, and finances. Another important circumstance for understanding sustainability is the extent to which a science program already was in place when the new science program originated. For example, a textbook program may have been in place for years prior to the arrival of the hands-on program. Or, it is possible that science simply wasn't taught at all at the elementary level and the hands-on program filled that void. Events as they relate to program origins are incidents that were actively pivotal in shaping the program and moving its development ahead to the point of establishment. For example, events could include a district's participation in leadership institutes, communication with science program leaders from other districts, securing grants and/or other funding, and opportunities for collaboration with partners. Together, these circumstances and events set a unique stage for the arrival of the hands-on program and, ultimately, influence its sustainability.

Longevity is simply the amount of time a program has been in place. It is tempting to hypothesize that programs in place for longer periods of time are likely to be more stable in the long term than younger programs. We have found that although all programs are to some extent vulnerable to district turbulence, this is the case. This stability might come from the assurance that elements of the program are accepted as the district standard, leaders' increased comfort levels, and widespread apparent success of the program. At the same time, however, longevity also can undermine a program because once a program has been in place for an extended amount of time, it is moved to the "back burner" of priorities. District leaders assume that the program is moving ahead, so they turn their attention to other areas, potentially leaving the science program to suffer from neglect.

3) Implementation and Adaptation

A range of implementation strategies have worked to create sustained programs. The establishment of some programs entailed a focus on all grades in a district simultaneously while others used a more gradual approach and focused on a sample of schools or particular grades. As the programs matured, other strategic decisions were made based on the varying circumstances and unique conditions. The strategic implementation choices in sustained programs were not driven by a commitment to a particular approach to reform. Rather, strategies that were successful demonstrated a sensitivity to and interaction with the district culture at that time. Many of these decisions were reserved for large-scale issues, such as how best to get the program off the ground, how to organize materials, and the best strategy for introducing a new component to the program. While these decisions were important in sustained programs, also important were the small-scale decisions that teacher leaders and program leaders make on a daily basis with regard to their interactions with individual teachers, colleagues, and administrators. These smaller but influential actions collectively provided significant support for the sustained program.

Hand in hand with the implementation strategies in the sustained programs studied were the strategic decisions to adapt the program. Some adaptations were proactive (e.g., applying for funds) while others were reactive (e.g., addressing the arrival of new state standards). Both required similar decisions and actions on the part of the program leaders and were critical contributors to the future of the sustained programs.

4) Perception
Perceptions—whether held by program leaders, program participants, or outsiders to the district—were significant supports and inhibitors of the sustained programs. In some cases, the perceptions of the programs differed greatly from the apparent actual status of the program. This was significant because, in the absence of enforced accountability measures, perception became a key driver of decision making for program adaptation and implementation. For example, the program leader may perceive that the program is at a particular level of implementation when, in fact, it is not. Or, the superintendent and other district administrators may perceive the program as being strong and exemplary. While this impression is positive, it also opens the door for potential neglect in allocations of future district dollars and attention.

5) Philosophy
Sustained programs directly and indirectly cultivated a widespread, shared philosophy. Philosophy in this context refers to beliefs held by a district's teachers, principals, and central administrators. We have identified two families of beliefs about science education significant in sustained programs: (1) whether science should be taught at all and (2) how science should be taught when it is taught. These two philosophical strands evolve, sometimes together, sometimes independently. In sustained programs, the second strand relating to how science should be taught was consistently strong—educators in these districts articulated beliefs that the hands-on approach to science instruction was the best way to teach science. However, the first strand representing belief in the importance of teaching science at the elementary level tended to fluctuate, depending on the changing district conditions. This suggests that a key contributor to sustained science programs is a strongly held belief in a particular pedagogy. However, even in the sustained programs in the study, the programs all still were vulnerable, as demonstrated by the variation of commitment to science instruction over time.

6) Quality
We defined the quality of a program as the extent to which its instruction and curricula facilitate positive attitudes toward, and student learning of, the elements of the scientific process and the basic concepts of the earth, physical, and life sciences. An assessment of the quality of the sustained programs in this study was beyond the scope of this research, and we do not have data sufficient to make substantiated claims about the quality of individual programs. However, we did identify two aspects of quality that have played an important role in sustainability. The first focuses on instruction and the absence of effective mechanisms for assessing the quality of instruction and the impact of professional development. The second relates closely to accountability for student learning. In the absence of such accountability measures, actual student learning of science concepts and processes had no impact on the sustainability of a program. When there were accountability measures, the program's quality was almost exclusively defined by evidence of student performance on that accountability measure. Thus, the alignment between the program and the district's accountability system become the primary indicator of program quality.

HIGHLIGHTS OF TWO FACTORS: ACCOUNTABILITY AND PHILOSOPHY
The 12 factors and 3 contextual issues summarized above are discussed in depth in the project's final cross-site report. This section highlights two of the factors with more detailed explanation and examples from the sustained programs in the study.

Accountability

The accountability issue, as explained above, falls into two main areas. First is accountability for student learning, and second is accountability for delivering the program. Generally speaking, standardized assessment in elementary science is rare. Three of the districts in this study reside in states that do test science at the elementary level and one has a district-level assessment, but the consequences of poor performance on those tests are tempered by the lack of high stakes and disproportionate attention given to the language arts and science test results. In cases where there is no test, teachers and principals in the districts often consider themselves accountable to the state standards and turn to them to guide their instructional choices.

The absence of mechanisms for ensuring that teachers are using the program was striking across the districts. We found that informal communications about the status of the program in any particular school were common and that judgments about actual implementation were mostly made based on the perceptions of the school and central office administrators. The one broad exception came in the case of using data from the materials management centers in the districts. Several program leaders looked at the extent to which materials were used when the kits were returned and interpreted that data to better understand how much of the program was actually being taught. Still, even with access to this data, program leaders seldom were in a position of power or authority that enabled them to act on their conclusions.

The accountability issue was one of the clearest examples of how any of these factors we identified can both support and inhibit sustainability of a program. Generally, many teachers, principals, and administrators felt that a call for accountability for student learning helped the program by providing it with a legitimacy and priority it didn't otherwise have. In the face of demands for information on student learning, however, the sustained programs found themselves vulnerable with a lack of concrete data to share when asked. Three sites—Lakeville, Fairfield, and Garden City—illustrate the complex nature of this factor and different ways it affected the sustained programs.

In Lakeville, there is no district- or state-level assessment in science at the elementary level. Standardized tests are administered at the high school, but the tests do not reflect the pedagogical approach promoted by the elementary program. The assessment methods proposed to teachers in the district include science lab notebooks and performance assessments, but they have not yet been developed. In the early 1990s, the district worked with an external consultant to develop embedded assessments, but at the end of the first year of work, the assessments were not completed and difficulties arose between the collaborating parties that prevented them from continuing. This left the program in a vulnerable position. In 2000, individuals in the community challenged the elementary science program by asking for a more traditional approach to instruction that more closely reflected the newly adopted state standards. In the absence of any sound, systematic data in support of the program, this sustained program faced the threat of termination. In a fortunate turn of events, the district was able to rally strong support from the teaching and public community and were able to combat the challenge and thus, preserve the program. But, the vulnerability of the program at that moment should not be underestimated. While having been sustained for some time, these programs remain fragile. Lakeville is taking steps to develop appropriate assessments and track student learning.

The impact of increased attention to accountability for student learning also is evident in Garden City. From the inception of the program until the time it was about eight years old, Garden City administered a criterion-referenced state test. That test had been in place for nearly 20 years but was dropped when the state created new standards. Recently, that test was replaced with another state standardized test that assesses language arts, mathematics, and science at the elementary as well as upper grade levels. With the long history of standardized tests, there is little resistance to the idea of a test. Quite the contrary, many feel that it will help to further the science program. They understand that the new test includes "critical thinking and open response" and reflects what they are trying to accomplish in the program. They feel the test will give the science program the clout it needs to convince principals to release their teachers for professional development and support the program. Indeed, this appears to be the case as evidenced in one principal's comment: "...the test is great...if we are going to value science...I can see it becoming a 4th R." Still, some are worried that teachers won't see the connection between the kit-based science program and the standardized test and, as a result, will drop the program in favor of more direct test preparatory instruction.

Teachers in Garden City and elsewhere could make this decision every day about the instruction they will use to support student learning. This illustrates the importance of the other aspect of accountability in sustained program: accountability for program delivery. Fairfield has a long-standing hands-on elementary science program with a history of more than 25 years. Fairfield has administered a number of standardized tests in science to its elementary students over the years, but program leaders recognized that because the test is not closely tied to the curriculum, its usefulness for interpreting the impact of their program is quite limited. In spite of Fairfield's long history and interest in having accountability measures for student learning, the fact remains that neither teachers nor principals are held accountable in any way for delivering the program to students. There are no formal mechanisms (such as mandatory observation of science instruction similar to the requirements for observing instruction in reading and mathematics), so the fate of the program rests with the individual commitment of the teachers and principals for providing their students with the science program. Under these conditions, the active involvement of the principal is critical to ensuring that teachers teach the science program. In an informal survey conducted as part of this study, a majority of principals reported that they supported science instruction while only half of their faculty felt their principal was an active supporter of the science program. The perceived lack of support on the part of these teachers, particularly when other subjects compete for classroom time, leaves teachers open to the possibility of foregoing teaching science. The implications for sustained programs is that the program itself can appear to be sustained—embedded in the system and accepted as standard practice—but not actually taught.

<u>Philosophy</u>

In the discussion of this factor, we have chosen four districts—Lakeville, Fairfield, Garden City, and Canton—because they each illustrate the importance of philosophy to adaptation and sustainability. Although the challenges facing these districts and their outcomes differed in scope and magnitude, the programs shared an important common characteristic that was consistent across the years. Each went through at least one adaptation of the program since its inception, and the programs that were in place following those adaptations subscribed to essentially the same philosophy as the programs that were in place prior to it. They were not distant cousins nor, arguably, close relatives, but at the core, they were the same kit-based programs that emphasized the primary importance of actively using science materials to teach the scientific process and science concepts. This firmly held belief in the importance of teaching science in this way served as the glue that held the programs together. To have an impact on sustainability, however, this belief needed to be vital in the minds of more than just the program leaders. It needed to be a motivator for teachers, administrators, and even parents and community members. The examples that follow will demonstrate how and why this was the case.

In Lakeville, after 15 years in place, the district's elementary science program was suddenly in serious jeopardy after a shift in political agendas on the school board created a public debate about its value. In response, program leaders assembled a public show of support for the program that spoke to its positive impact and its ability to meet the standards for science that had been accepted by the district. This response included statements from teachers, parents, and community members of all kinds in support of what the science program had achieved for the district's students. The leaders' efforts successfully focused on managing the response to scrutiny and criticism, and the result was that the program survived intact. For this strategy to succeed, however, a deep understanding of the science program as well as support for it needed to be widespread among the public. Without this shared understanding of the values of the program, it would have been impossible to assemble such a public outpouring of support. Moreover, the testimony that the public brought to bear was powerful not only because it was broad-based, but also because it was rooted in experience over time. The evidence submitted by supporters was not the result of hastily coached volunteers, but of those with personal experience who had seen what the program had achieved since it began. In Lakeville, belief in the hands-on philosophy was widespread and had grown over time.

In Fairfield, a regular adoption cycle occurred at about the 10th year of the program and prompted an examination of the kits and the curriculum. This examination revealed the need to update the kits, and the leaders' response was to undergo a large-scale revision process that required the time and attention of many teachers over a two-year period. The revision produced re-designed kits with, among other things, more carefully conceived activities, targeted teacher guides, unit objectives, and student assessments. In this case, the district's commitment to the program and the science kits made the revision process the natural response to this regularly scheduled examination of curriculum. However, it is important to recognize that the reason the process was suggested by the program's leaders in the first place, and successful in the second, was because the commitment was embedded at all levels of the district from the classroom teacher to the program leader to the central administration. This commitment enabled the revision to take place and explains the resulting strengthened science program.

In Garden City, the influence of the state standards in science brought on an examination of the science program's content during the program's maturation phase. To ensure that all areas included in the state standards were covered in the curriculum, the program's leader implemented adaptations to the program's curriculum. These adaptations included changing the grades in which specific topics were taught, buying new kits, and including the use of textbooks to address topics that were not covered by kits. Garden City may be the district where the most significant alteration to the program took place because of the introduction of textbooks to the curriculum. However, rather than suggest a weakening of commitment, the response to textbooks appears to be evidence of the value that teachers place on the use of the kits. In our discussions with teachers, the possibility that textbooks would supplant the kits was not a concern. Covering all of the material in the standards was the challenge, compelled by the close watch that was kept on all subject areas' reliance on them. As was shown in Fairfield, revising the kits to address the standards would have been a lengthy process and would not have been a viable option in this case; the pressure to respond to the standards was too great and required an immediate solution. Maintaining the emphasis on the kits, beginning the process of updating them by making new purchases, and using textbooks to fill in gaps, was evidence of and reinforcement for the importance of the use of materials to teach science. Here again, the response required guidance by the program leader. But that alone would not have ensured success; it was the shared commitment across the district that made this adaptation succeed.

Finally, Canton presents the most extreme example of adaptability and endurance. Here, a powerful, political shift away from a centrally governed school district to one in which control resided within each school eliminated all centralized support for curriculum and instruction. After more than 30 years of strong development and evolution, the science program (as well as other subject areas) was dismantled. The position of science program leader was eliminated, along with instructional support for teachers and support for the supply of materials, which is so critical to a kit-based program. As a result, by most definitions, the program disappeared and, in fact, it became invisible at the district level. However, our interviews revealed that rather than disappear, the program went "underground," and those teachers committed to science continued to use the kits and provide science instruction to their students. Although program leaders could not adapt to the change in circumstance, many of Canton's teachers certainly did, as they found ways to continue to use the kits to teach science. Teachers and administrators who were there throughout this period refer to it as the program's "dark time," but they do not suggest that it died. As the need for centralized functions subsequently became apparent and a centralized science program became acceptable to the community, the science program re-emerged, again using a kit-based approach. In fact, as the new curriculum was being planned, the use of kits was *assumed* and a traditional textbook program was not considered. It would be misleading to suggest that, during the "dark time," all teachers continued to teach science or that all of those who did teach science used the kits. However, enough of them did use kits to sustain the institutional memory of the importance of the hands-on approach. Although considerable effort was required to re-establish hands-on science after many years of neglect, the program's new leaders did not start from scratch. A districtwide, elevated understanding of kit-based programs formed a foundation upon which the new program was built. Canton is an example of how powerful the understanding and acceptance of the hands-on approach is when it is in the hearts and minds of teachers and principals. Leaders can use these values to guide their strategic decisions, but without the commitment from those who ultimately deliver the program, the direction they provide will not be enough to move the program ahead.

SUMMARY

School districts are complex and dynamic environments requiring that reforms stay focused on concrete goals while adapting to changing conditions. Educators require more knowledge about how to maintain their programs, strategically concentrate their efforts, and build capacity for continuous growth in such challenging arenas. But there is no one formula for sustainability. This study confirmed and documented the complexities that school districts face when engaged in such endeavors, identified factors that contribute to and inhibit sustained science education programs, and recorded the diagnostic and decision-making processes leaders use to advance their programs.

Funders have invested millions of dollars of seed money with the hope that this funding, if wisely utilized, can leverage durable, large-scale changes. However, in spite of these resources and educators' best efforts, improved science education programs still have tenuous futures. If widespread, sustained change is not typically the result of these investments, then the rationale for making this type of large-scale reform effort is questionable. It is critically important for educators to draw from the findings of this study and other works that identify lessons from past efforts to inform the design, implementation, and adaptation of new reforms. The findings of this study include practical suggestions as well as theoretical ideas, both of which are important for science education leaders to consider throughout the life of their programs, from establishment, growth and maturation, and finally to sustained evolution.

Table I: Overview of School Districts

	GLENWOOD*	LAKEVILLE	HUDSON	MONTVIEW	NORWAY	GARDEN CITY	SYCAMORE	BENTON	BOLTON
SIZE									
Sq. Miles	47	76	200	800	**	800	25	15	320
# elem. students	27,000	12,000	43,151	47,087	5,849	28,000	6,400	4,600	28,000
# elem. schools	77	23	50	92	23	52	30	14	60
# elem. classroom teachers	1,300	778	1,630	1,978	600	1,300	300	200	1,144
RESOURCES									
Per pupil expenditure	5,668	4,996	5,122	3,198	5,973	5,659	6,500	17,000	6,508
Teacher starting salary	$31,172	$35,573	$27,686	$25,832	$27,467	$27,718	$29,892	$34,116	$32,600
NSF funds?	yes	yes	yes	no	no	no	no	yes	yes
DEMOGRAPHICS									
% students eligible for free and reduced price lunch	66%	70%	41.1%	58%	40%	32%	65%	39%	30%
% white	13	17	68	85	57	69	69	41	62
% African American	18	34	3	1	12	28	12	34	9
% Hispanic	21	45	23	10	10	0	11	14	6
% Asian/Pacific Islander	27 (Chinese)	4	2	3	18	0	8	10	9
% Native American	21	0	4	1	3	0	0	0	13
% Other	0	0	0	0	0	3	0	1	1
OTHER INFORMATION									
Year program began	1989	1986	1974	1968	1966	1989	1988	1994	1974

* District names are pseudonyms.
** Unable to obtain information from school district.

REFERENCES

Covaleski, J. (1994). The educational system and resistance to reform: The limits of policy. *Education Policy Analysis Archives, 2* (4).

DeBoer, G. (1991). *A history of ideas in science education*. Teachers College Press: New York, NY.

Loucks, S. & Hall, G. (1977). Assessing and facilitating the implementation of innovations: A New Approach. *Educational Technology, 17* (1), 18–21.

McCormack, A. (1992). Trends and issues in science curriculum. In *A science curriculum resource handbook: A practical guide for K–12 science curriculum* (pp. 16–41). Kraus International Publishers: Millwood, NY.

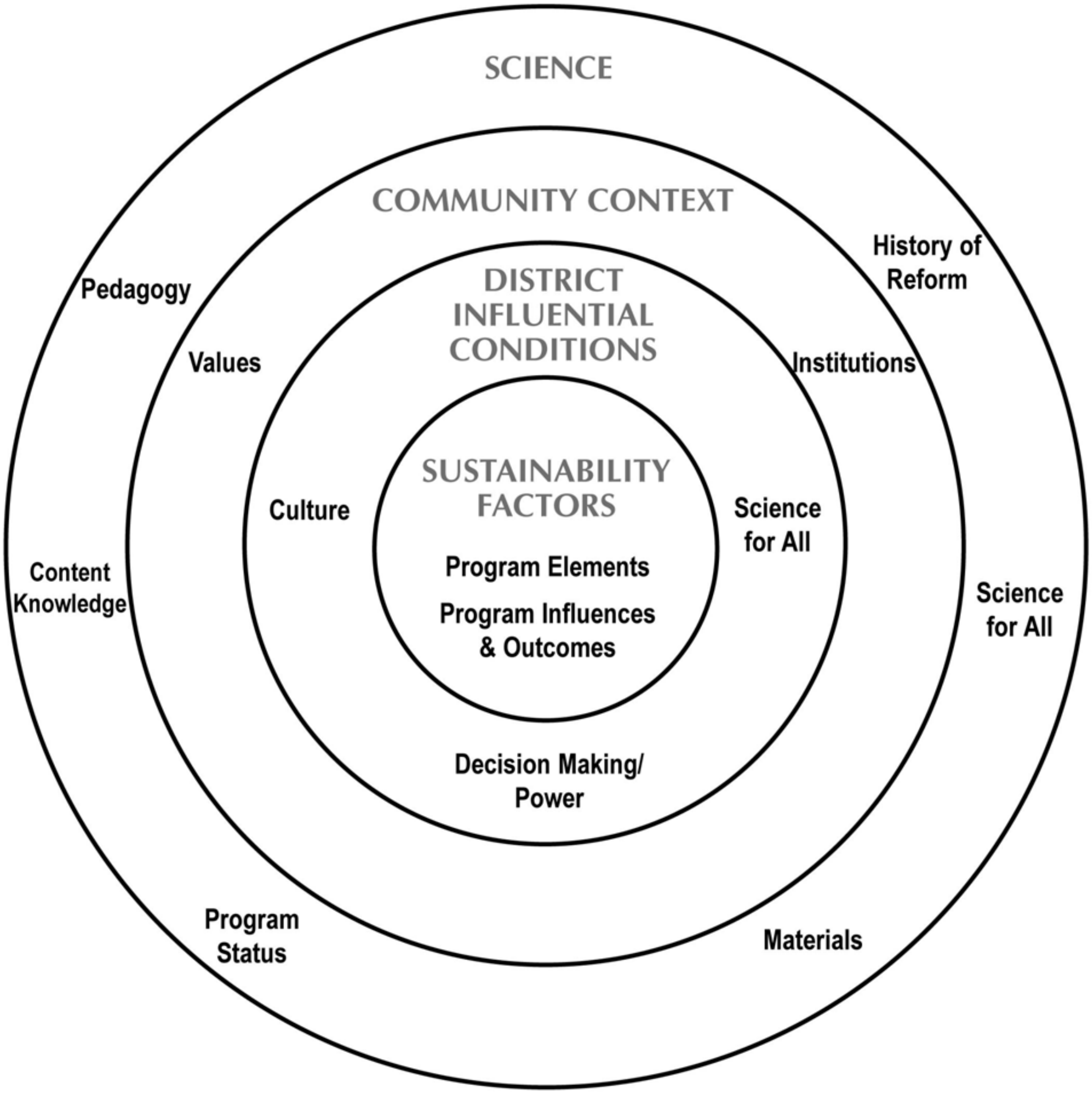
McLaughlin, M. (1989, March). The RAND change agent study ten years later: Macro perspectives and micro realities. Paper based on address given at the Annual Meeting of the AERA (San Francisco, CA, March 1989). Washington, DC: OERI. (ERIC document ED 342 085).

Portelli, J. (1987). Perspectives and imperatives on defining curriculum, *Journal of Curriculum and Supervision, 2* (4), 54–67.

Sarason, S. (1996). Revisiting the culture of the school and the problem of change. Teachers College Press: New York, NY.

Sashkin, M. & Egermeier, J. (1992). School change models and processes: A review and synthesis of research and practice. (OERI, United States Department of Education Working Paper, PIP 92-9).

APPENDIX


SCIENCE
COMMUNITY CONTEXT
DISTRICT INFLUENTIAL CONDITIONS
SUSTAINABILITY FACTORS
Program Elements
Program Influences & Outcomes
Pedagogy
History of Reform
Values
Institutions
Culture
Science for All
Content Knowledge
Science for All
Decision Making/ Power
Program Status
Materials