CROSS-SITE REPORT

RESEARCHING THE SUSTAINABILITY OF REFORM

FACTORS THAT CONTRIBUTE TO OR INHIBIT PROGRAM ENDURANCE

JEANNE ROSE CENTURY
ABIGAIL JURIST LEVY

CENTER FOR SCIENCE EDUCATION (CSE)
EDUCATION DEVELOPMENT CENTER, INC. (EDC)
NEWTON, MASS.

NATIONAL SCIENCE FOUNDATION

ACKNOWLEDGMENTS

This cross-site analysis is the result of research in nine school districts. We would like to thank the site leaders in each of the study's sites for their support, hard work, and frankness throughout the data collection process. We are grateful to the teachers, principals, district administrators, and many others who spoke with us in each site. Additionally, we would like to acknowledge the efforts of the staff at the Caltech Pre-College Science Initiative (CAPSI), and particularly Jerry Pine, for their input and feedback on this report. We also would like to acknowledge the role that EDC staff member Felisa Tibbitts played in the development of this report. Further, we would like to acknowledge the insights of members of the RSR advisory board and the thoughtful attention and support of EDC staff members Judi Sandler and Karen Worth. We also are grateful to Kerry Ouellet for her tireless efforts in the editing and layout process.

Center for Science Education
Education Development Center, Inc.
55 Chapel Street
Newton, MA 02458-1060
800-225-4276

This report and other project information can be found at http://www.edc.org/cse

This material is based upon work supported by the National Science Foundation (NSF) under Grant No. REC-9805078.

CROSS-SITE REPORT

TABLE OF CONTENTS

PROLOGUE

Over the past decade, extraordinary resources have been committed to promoting the improvement of science education. Since 1990, a conservative estimate suggests that the National Science Foundation (NSF) has invested well over $450 million to improve K–12 science education through comprehensive reform efforts. If one also includes NSF's investments in curricula and materials, that amount would significantly increase. Additionally, many school districts have invested their own resources along with countless hours in the planning and implementation of NSF's investments in their schools. And still other districts have simply taken on the science education reform endeavor at their own expense. These seasoned educators at all levels of the education system have seen reforms come and go due to shifts in the political tide, changes in school and district leadership, or increases in the popularity of different pedagogical approaches. So, if there is one question ever-present on the minds of teachers, administrators, funders, and policy makers working to improve education for their students, it is: *How do we ensure that the programs we are implementing will last?*

The reform efforts referred to above have centered on establishing science programs that promote the use of hands-on materials through an inquiry approach. Creating such a program, or changing a districtwide science program from one that is driven by textbooks to one that is centered on using materials, is a formidable task with far reaching implications. District budgets, teacher training and professional development, articulation and alignment, testing and assessment processes, curriculum design, and science standards and frameworks are only some of the districts' policies and practices that program leaders and administrators need to address. Without sufficient planning, resources, resilience, and fortitude, initial investments are unlikely to bear fruit, no matter how dedicated the leaders. And even with a full complement of all of the above, educators ask, how long should it take for a hands-on program to become embedded in a district's culture? And what else does it take to make sure that the initial efforts will ultimately pay off in the form of an enduring, hands-on, districtwide science program?

Indeed, these questions were prominent in the minds of the practitioner colleagues of staff from the Center for Science Education (CSE) at Education Development Center, Inc., (EDC) in Newton, Mass., and the Caltech Pre-College Science Initiative (CAPSI) at the California Institute of Technology in Pasadena, Calif. CSE and CAPSI both had a strong foundation of working with school districts to plan and implement districtwide, hands-on elementary science education programs and had concerns about the sustainability of their fledgling programs. Moreover, CSE and CAPSI staff had seen the price that districts paid when their reform efforts failed. Resources were clearly lost, but so was time, effort, morale, opportunities to expand teachers' content knowledge and teaching skills, not to mention opportunities for children to experience science in ways that heretofore had been unavailable to them.

Thus, motivated by the importance and urgency of these sustainability questions, CSE and CAPSI collaborated to obtain funding from NSF for a research effort aimed at answering the question: *What contributes to or inhibits the sustainability of a districtwide, hands-on inquiry science program?* The project, which came to be known as Researching the Sustainability of Reform (REC-9805078), attempted to answer this question through a three-year study of nine districts in the United States that had districtwide, hands-on inquiry science programs in place from nearly 10 to 30 years. Findings from this research are contained in nine site-specific reports and in this cross-site analysis, which discusses the broader findings, trends, and themes gleaned from all sites.

This cross-site analysis is organized into four parts. Part I provides an overview of the study including design and methodology. Part II sets the stage for understanding the findings by presenting some of the overarching ideas that emerged from the study and providing the reader with a concrete portrayal of what these programs look like, how they develop, and the ways they are implemented. Part III is a discussion of the findings themselves including the contexts and conditions that influence sustainability, factors that pertain to the individual elements of a science program, and factors that affect the program as a whole. Finally, Part IV focuses on the implications of the findings for leaders of individual science programs and for the field.

Background of EDC's Center for Science Education (CSE)

CSE's history of working with school districts to improve their science education programs is grounded in 15 years of curriculum development, professional development, technical assistance, and research, all conducted in collaboration with practitioners in urban, suburban, and rural school districts. In 1987, CSE began its work focusing on inquiry-based science curricula by developing *Insights: A Hands-On Elementary Inquiry-Based Science Curriculum.*[1] Not long after, CSE developed *Insights* for the middle level, and then *Insights in Biology* for grades 9 and 10. CSE staff now are developing materials for the Pre–K, elementary, middle, and high school levels.

Concurrent with curriculum development work, CSE has provided technical assistance and professional development support to over 300 districts across the country, many of which have NSF-supported teacher enhancement and systemic reform projects. Much of the work has been in close collaboration with science directors/coordinators of district school systems

---

[1] *Insights: An Elementary Hands-On Inquiry Science Curriculum.* Developed by Education Development Center, Inc., 1997. Published by Kendall/Hunt Publishing Company.

as well as superintendents, assistant superintendents, principals, and teacher leaders. CSE has conducted seminars and institutes on many issues related to science education reform, including assessment, science and literacy, science standards, and increasingly diverse student populations. In addition, CSE staff, along with leaders of several mature districtwide, hands-on science programs, provided direct technical assistance to districts beginning their work in science education reform. Part of this work included the development of materials and resources for leaders of these fledgling programs, including the first monograph in the NSF *Foundations* series[2].

Using this foundation of knowledge and experience, CSE also has developed a body of research and evaluation work grounded in a commitment to conducting rigorous studies that provide useful, practical information to educators engaged in education reform. The research work is a natural outgrowth from the Center's curriculum writing, professional development, and technical assistance efforts that regularly raise many research issues and evaluation questions—in this case, the focus is on sustaining reform. The research and evaluation work includes a range of methodologies, purposes, and approaches. It reflects the beliefs that research studies should result in findings that are directly applicable in the field; research questions should emerge from field-based experience and issues of direct importance to practitioners; and evaluations should provide information of practical and immediate use to the client.

Thus, this research project was a natural fit for CSE and its practitioner colleagues. CSE staff understand that program leaders' abilities to make the case for inquiry-based or hands-on science education, guide materials selection, develop professional development programs, and provide overall leadership make progress possible. But still, it cannot completely guard against their programs' vulnerability to the shifting pressures that accompany political and community change. This research project sought to reduce that vulnerability with understandings and strategies identified by studying those places that had found a way to survive.

For more information about CSE, visit the CSE Web site at http://www.edc.org/CSE. For more information about EDC and its other areas of work, visit http://www.edc.org.

Background of CAPSI

CAPSI (Caltech Precollege Science Initiative) was founded in 1985 as a collaborative effort of Caltech scientists and the Pasadena Unified School District to initiate a K–6 program of hands-on inquiry science in the schools. Begun on a small scale with volunteers in one school, the program

---

[2] National Science Foundation. (1997). *Foundations: The Challenge and Promise of K–8 Elementary Science Reform.*

was expanded to the entire district of over 10,000 students in 23 elementary schools. This scale-up, with NSF support, became a model for the NSF Local Systemic Change Initiative. In the 1990s, many educators from across the United States and from overseas visited the program, observed classes, and consulted with the leaders on how to implement their own programs, many of which became successful districtwide efforts. In addition, scientists and engineers in France, Estonia, and Colombia have built on the CAPSI model of collaboration with educators to begin to implement national programs in their own countries.

CAPSI expanded its activities into the development of both pre-service and in-service science content courses for elementary teachers in the late 1990s, which have been successful in a variety of school districts across the nation and in Los Angeles area colleges. At that time, CAPSI and the Pasadena Schools collaborated to apply for and win the first NSF Center grant for teacher enhancement, to work with 14 predominantly minority school districts in California in establishing inquiry-based K–6 programs. After seven years, the Center still supports the continuing growth and development of 10 districts, which have formed a unique closely-knit consortium of K–6 reformers.

CAPSI's experience helping to establish districtwide science programs and coping with the problems of sustainability matched the experience of the leaders in the Center for Science Education at EDC. Together, they proposed this study on issues related to the sustainability of K–6 inquiry science programs. This initial research effort by CAPSI has grown to encompass a variety of other studies, all closely related to the practice of inquiry science education. These include a comparative study of fifth graders' science abilities in hands-on and textbook-based programs; a study of an Internet-based interactive site that appeals particularly to middle school girls; and a study of how best to use science notebooks in K–6 classrooms and their impact on science and literacy learning. In addition to the work of the Center and the Research Group, CAPSI has embarked on a project to develop next-generation inquiry curricula for grades 7–10, with field-tests of the first units beginning in 2002–03.

CAPSI has been identified by the National Academy as an exemplar of scientist-educator collaboration, and is featured on their Web site at www.nas.edu/rise/examp81.htm, while the CAPSI Web site is at www.capsi.caltech.edu.

PROJECT DESCRIPTION

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 to 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. This research was supported by a grant from the National Science Foundation and was directed by Dr. Jeanne Rose Century at EDC and Dr. Jerome Pine at CAPSI.

The goal of this study was to identify and document factors in school systems that contribute to sustained educational change in science education. The purpose was to provide districts now engaged in improving their science education programs and districts that are considering doing so in the future with information to 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.

Through on-site interviews and observations, surveys, case studies, and document analysis, the study investigated the districts' efforts in the following areas:

- Current status of the science program compared with initial goals
- System context and external conditions that have an impact on lasting change
- Strategies for achieving program goals and building district capacity to improve
- The influence of practitioner and system capacity on sustainability
- External funds as a catalyst for widespread, lasting reform

The findings of the research include nine descriptive site summaries and a cross-site report. The site summaries were designed primarily to provide the reader with a description of the origins, implementation, and evolution of each of the nine science programs. They also offer a brief analytic section that is designed to provide the reader with a bridge to the cross-site report. 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 drawn from the nine sites. The cross-site report also includes a discussion of implications of the findings for funders, reformers, and practitioners.

Please direct any inquiries about this study to:
EDC Center for Science Education
55 Chapel Street
Newton, MA 02458
617-969-7100

Dr. Jeanne Rose Century	Abigail Jurist Levy
x2414	x2437
jcentury@edc.org	alevy@edc.org

To download site reports from this study, visit the CSE Web site at http://www.edc.org/cse

OVERVIEW OF PROJECT SITES

	GLENWOOD[*]	LAKEVILLE	HUDSON[††]	MONTVIEW[‡]	BAYVIEW	GARDEN CITY	SYCAMORE	BENTON	BOLTON
SIZE									
Sq. Miles	47[†]	76	200	800	55	800	25	15	320
# elem. students	27,000	12,000	43,151	47,087	5,849	28,000	6,400	4,300	27,000
# elem. schools	77	23	50	92	23	52	30	15	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	4,443	5,973	5,046	6,500	13,296	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%	18%	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	11	10	0	11	14	6
% Asian/Pacific Islander	27 (Chinese)	4	2	3	18	0	8	10	9
% Native American	21	0	4	0	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	1977

[*] District names are pseudonyms.

[†] Figures are for years ranging from 1998–2000. During this time demographics and expenditures shifted and were calculated in a variety of ways.

[††] The Hudson site report offers the reader an additional detailed description of a classroom science lesson.

[‡] The Montview site report is unique in that it emphasizes the historical development of the program and the circumstances that influenced and shaped its evolution.

CROSS-SITE REPORT

EXECUTIVE SUMMARY

INTRODUCTION

Since 1990, a conservative estimate suggests that the National Science Foundation (NSF) has invested well over $450 million to improve K–12 science education through comprehensive reform efforts. Additionally, many school districts have invested their own resources along with countless hours in the planning and implementation of NSF's investments in their schools. And still other districts have simply taken on the science education reform endeavor at their own expense. Educators at all levels of the education system have seen reforms come and go, and one question ever-present on their minds is: *How do we ensure that the programs we are implementing will last?*

Motivated by the importance and urgency of these sustainability questions, the Center for Science Education, Newton, Mass., and California Pre-Science Initiative, Pasadena, Calif., collaborated on a research effort aimed at answering the question: What contributes to or inhibits the sustainability of a districtwide, hands-on inquiry science program? The project (REC-9805078) addressed this question through a three-year study of nine districts in the United States that had districtwide, hands-on inquiry science programs in place from nearly 10 to 30 years. Findings from this research are contained in nine site-specific reports and in this cross-site analysis, which discusses the broader findings, trends, and themes gleaned from all sites.

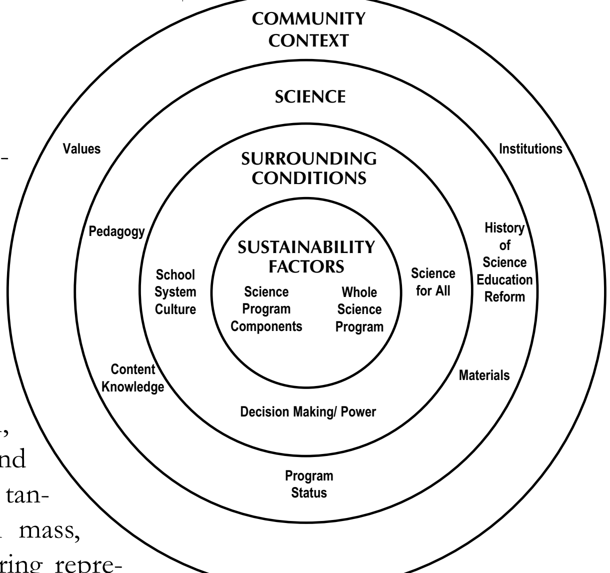
SETTING THE STAGE FOR THE FINDINGS

MAPPING THE FINDINGS

Sustainability as it relates to the findings of this study is complex and covers a spectrum of factors. Figure 1, Sustainability Factors and Surrounding Conditions, maps the factors and conditions important to sustainability and represents the complexity of their interactions with one another.

In the center are the most concrete factors that contribute to or inhibit sustainability—those that pertain to specific science program components (accountability, implementation, instructional materials, leadership, money, partnerships, and professional development) and those that are somewhat less tangible and pertain to the whole science program (critical mass, adaptation, perception, philosophy, and quality). The next ring represents factors (district culture, decision making and power, and science for all) that influence the conditions that have some bearing on the operation

Figure 1
Sustainability Factors and Surrounding Conditions



of the program in the district context and on the strategies program leaders employ to support the program's stability and development. These factors are the core findings of this study.

The next ring represents science, itself, and the array of unique issues it brings, particularly when implemented at the elementary level. And finally, the broadest ring in the diagram represents the community context. This refers to the values and institutions that predominate, influencing the program and shaping program leaders' decisions. Although the community context is, perhaps, the most removed from an elementary science program's daily work, it clearly exerts pressures that can play a powerful role in a program's constancy and growth.

DEFINING SUSTAINABILITY AND THE PHASES OF CHANGE

To coherently understand this study's findings, the reader must first understand two important concepts: what we mean by sustainability and the ways in which science programs experience changes over time.

Defining Sustainability—Maintenance vs. Sustainability

> *Sustainability: The ability of a program to maintain its core beliefs and values and use them to guide program adaptations to changes and pressures over time.*

Educators commonly view sustainability as program maintenance—embedding a program, as designed, into a standing operating system. By this definition, anything short of a replica is not sustainability.

This project found that "sustaining districtwide education reform" is 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 this research as researchers sought the answers to two questions: (1) Was the program 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 research progressed, it became clear that none of the programs were exact replicas of their earliest years, and the longer the time horizon, the more clearly the trends in evolution emerged. Thus, it was important to make a clear distinction between program maintenance and sustainability. A program is maintained if its basic elements are well established and commonly accepted as standard practice. Sustainability, on the other hand, stresses the importance of adapting and improving in response to the changes that inevitably occur in a school district. A program must be maintained before it can reach sustainability, but it cannot be stalled at maintenance; it must develop an ability to evolve and adapt. But adaptability alone is not enough; adaptations must be guided by the essential values and beliefs that characterize the core of the program's intent. It is the continued influence of those beliefs and values that ensures that, as programs evolve, they remain closely connected to their earlier generations.

Three Phases—Moving Toward Sustainability Over Time

The research identified three stages of program development that advance programs from maintenance to sustainability: *establishment, maturation,* and *evolution*. The lines of demarcation between phases are not exact; and programs do not always move forward smoothly. They may advance, hold, slide back, retrench, and then move ahead again. But, the longer a program's time horizon, the more clear its pattern of growth and development.

The *establishment* phase focuses on the very concrete elements of the program, making sure that they are well established, accepted, and working efficiently and predictably districtwide.

The next developmental phase is *maturation*. Here, the focus is 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.

The third phase of development is *evolution*. The hallmarks of the evolution phase are growth and improvement.

Programs never shed entirely the threats and challenges of earlier phases. Rather, leaders continue to address ongoing issues as they take on a new set of goals associated with their continuing development. Moreover, with each additional set of goals, there are important implications at all levels of a school system: the classroom, the school, and the district. To be sustained, program goals must be realized at different levels, which require 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 the factors identified in this study at any of these different levels of the system. Together, the phase of development and the program leaders' level of orientation determine the factors' importance and priority.

FINDINGS

INTRODUCTION

The stories of the elementary science programs in this study are complex. Many factors have contributed to and inhibited their sustainability over time. These factors do not operate in isolation; they interact with each other, shift in importance and influence over time, and are often difficult to distinguish from one another. To discuss them, it is necessary to draw somewhat arbitrary distinctions between them, but their web-like relationships are a finding in and of itself. It explicates the range of pressures that come to bear on the sustainability of a program and the difficulty program leaders face in anticipating or controlling for them.

SECTION 1: FACTORS THAT PERTAIN TO SURROUNDING CONDITIONS

School System Culture

- *A shared culture of collaboration and respect can support the establishment, growth, and evolution of sustained programs, while a competitive culture that illuminates rivalries can inhibit them.*
- *Even when there is individual will and interest, a district culture that lacks established communication avenues can stand in the way of taking actions to support a sustained program.*
- *Tensions between centralized services and a decentralized district culture can negatively affect sustained programs.*
- *A district culture that promotes learning and outreach can benefit sustained programs.*

In this project, culture refers to the nature of the human, structural, and systemic environment in which science programs 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 expected practices (e.g., volunteerism, support for professional growth, and extent of support for innovation).

Though project leaders might define culture in a far less formal way, they do not underestimate its power. They work within its confines and recognize that it affects the ways their goals, strategies, and communications are interpreted. As an influential condition, culture sets a foundation for the ways and extent to which the other factors described below contribute to and inhibit the sustained program. Thus, efforts to bring a science program to fruition must be compatible with the culture or, even though well intentioned, they are likely to fail.

Culture is a backdrop that influences program leaders actions and the interactions of the factors that support and inhibit sustainability. Readers will be able to draw links between culture and some of those factors, including decision making, leadership, implementation, money, and adaptation (to name a few) with ease. Simply put, culture is pervasive and, though at times difficult to accurately describe or interpret, a key influence on the district operations surrounding the establishment, growth, and evolution of the hands-on programs in this study.

Decision Making and Power

- *Program leaders have little formal decision-making power or authority over the elements of their science programs.*
- *Decisions are made at many levels in a district by many different stakeholders. Any single decision can advance or inhibit the status of the science program.*
- *Leaders of sustained programs must find ways to navigate the decision-making structures in their districts and gain access to those who have the power to influence the status of their programs.*
- *The support of the central office is critical to the well-being of the science programs.*

Leaders have relatively little control over the many pressures and issues that can and do influence the growth and development of districtwide hands-on science programs. In fact, there is a wide range of decision makers in a district, each with his or her own allotment of formal and informal power, who can advance or inhibit a science program's growth and development at any point in time. Given this relative lack of control, leaders of sustained programs must understand their district's power structure and be adept at negotiating it in order to exercise what influence they can over the decisions that will affect their programs.

Each of the districts in this study has its own style and process for making policy and program decisions, some more explicit than others. These different styles and allocations of power form the landscape within which science program leaders try to advance their programs. Also important in the program leader's landscape are the many different levels, district, school, and classroom, at which decisions are made about whether and how the science program will be implemented. Clearly, gaining access to those with power and decision-making authority is key.

Science For All

- *A centralized (or districtwide) program is considered an equalizer for schools and students, who may otherwise experience inequitable distributions of resources and variable classroom experiences.*
- *In the absence of accountability, equity suffers.*
- *Given the equalizing nature of a districtwide science program, when equity is expressed as a goal and value of the district, that goal isn't necessarily translated to support for the science program.*

"Science for All" 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, the issue of equity emerged as a factor in three main areas: access to science instruction, equitable implementation of the program, and the value of the science program for specific populations.

Teachers and administrators recognized the science program as an equalizer with regard to materials and curriculum, due in part to schools' widely varying levels of economic support and inconsistencies in curriculum in other subject areas. Moreover, all of the districts in the study have systems that ensure all schools have access to the science program; in fact, this is one of the features that defines the programs as districtwide. Yet, despite leaders' best efforts, the data show that program use within each district has been highly variable. The study found no evidence in any site of a districtwide system in place to assess whether or not teachers are actually teaching science, and there are no districtwide consequences for teachers who fail to do so. The end result, then, is that instruction is left to the discretion of the teacher, resulting in inconsistent and, by definition, inequitable instruction. Though many recognized the districtwide science program's potential, not only to provide science instruction to all students but also to contribute to making progress toward improved equity across the district, the interest in supporting this potential was never clearly articulated either verbally or in writing in any of the data collected.

SECTION 2: FACTORS THAT PERTAIN TO SCIENCE PROGRAM COMPONENTS

Accountability

- *There is limited accountability for student learning or for the delivery of the program. This can either contribute to or inhibit the sustainability of the science program depending on the district context.*
- *In the presence of high visibility and high stakes tests, science is often overshadowed and, therefore, time and resources devoted to its accountability are diminished.*
- *When an accountability strategy for student learning or program delivery does exist, resulting data are of little use to program leaders if they have no power or authority to make and follow through on decisions based on that data.*

Two types of accountability have played a role in the sustainability of the hands-on science programs in this study: accountability for student learning and accountability for principals' and teachers' program delivery. Accountability measures for student learning include student written and performance tests, student work, and writing in student science notebooks. Accountability measures for program delivery, on the other hand, include requirements for principal observations of science instruction, tracking of kit usage, and analysis of school improvement plans. Generally speaking, some districts have district or state tests in place that provide the only mechanism for accountability for student learning. Mechanisms supporting program delivery, however, are universally weak. The presence and absence of these mechanisms, depending on the site and its context, sometimes support the sustained program and sometimes hinder it, but always cause high levels of concern and anxiety.

In the face of demands for information on student learning, sustained programs can be vulnerable. At the same time, however, most of the programs in the study thrived for many years with no such data. This suggests that, in the absence of specific accountability measures, program leaders and others make decisions based on limited and informal data sources combined with their own observations and perceptions about the status of the program. Thus, a program can appear to be sustained—embedded in the system and accepted as standard practice—but not actually taught.

Implementation

- *Leaders of sustained programs have used a range of approaches to implementation with no single approach demonstrating more success than another.*
- *Central office support is a necessity for laying the groundwork and establishing the elements of a sustained program.*
- *Leaders of sustained programs choose implementation strategies that account for the culture of the district, district priorities, and the relative importance of the different elements of the program at a given time.*

Implementation refers to the strategies program leaders use to initiate hands-on science programs and the methods they use to bring their science programs to be accepted as districtwide practice. Though all of the district leaders in the study have shared a similar challenge—establishing a program that includes resources, curriculum, professional development, and instructional materials—their overall approaches to implementing their programs have been highly variable. It is worth noting that although each leader could have chosen to pursue any kind of science program, each chose to focus on hands-on science instruction. Whether their belief in the hands-on approach has come from exposure via a mentor or colleague, personal experience with hands-on instruction, or their own science background, all have been deeply committed to bringing the hands-on experience to their communities. Programs also were influenced to some extent by the national political climate of the 1960s and 1970s that followed Sputnik and was concurrent with NSF's emphasis on developing science curricula and increasing the number of people pursuing careers as scientists. Given the range of strategies that has worked for the districts in this study, one can conclude that no single approach to implementation necessarily leads to a sustained program.

Instructional Materials

- *The curricula of sustained programs typically are composed of a combination of materials—ranging from homemade lessons to commercial units—and often have supplemental components which, in some cases, include textbooks.*
- *Instructional materials in sustained programs evolve and are adapted over time.*
- *A district materials management center provides symbolic and practical evidence that a hands-on science program has been sustained.*
- *Instructional materials for hands-on elementary science programs require processes and systems for development and selection; management, distribution, and storage; and acquisition and refurbishment that consume a great deal of human and financial resources.*

Instructional materials are an essential component of any science education program. All of the science programs in this study were primarily kit-based, meaning they were based on boxes that included a teacher's guide and the necessary manipulatives for teaching the lessons outlined in that guide. From the very start, program leaders in every site had a shared challenge—what materials to use; how to get those materials to the teachers; and subsequently, how to retrieve them and prepare them for the next teacher. While sharing similar concerns, they each devised a sensible, customized strategy given the financial resources, climates, and cultures of their districts.

Not only are materials centers necessary, practical supports for the science programs, but they also make an important symbolic contribution to the programs' sustainability. In some districts, the centers are viewed as a point of pride and perceived, to some extent, as evidence that the district is giving attention and support to elementary science instruction. Thus, eliminating the materials center would be tantamount to cutting the program. As a result, one can speculate that other areas of the program that are equally important but less visible and concrete (e.g., professional development) are targets instead.

Leadership

- *The requirements of a sustained program's leadership vary at different stages of the program and with shifting district conditions.*
- *The style of leadership needs to coincide with the culture of the community and the needs of the program.*
- *Attempts to develop the engagement of school-level leaders have largely been unsuccessful.*
- *Superintendents have three tools they can choose to exercise or not: authority, political influence, and budgetary influence.*
- *Program leaders and their leadership teams are ambivalent about the more supervisory and coaching roles they might play.*

Leadership in sustained programs is wide ranging and evident at all levels of the system. It extends from formally identified leaders to informal or "behind the scenes" leaders. Leaders of the programs in this study have had widely varied strengths and weaknesses, but their ultimate success has been dependent on their abilities to be flexible, respond to shifting district conditions, and interact appropriately with the local culture. Their experiences have offered insights into how leaders at all levels in a district can contribute to sustained programs.

Program leaders all have been intelligent and passionate about their work, with the management skills to enable them to realize their visions, albeit with different styles and approaches. Different leadership skills are required for the various stages of program development—establishment, maturation, and evolution—and although, generally speaking, the tasks remain consistent from place to place, each district's culture and operating systems require different strategies to accomplish them.

Another key leadership influence rests with the superintendent, who can exercise power over the budget, accountability measures, and political relationships. In addition to reaching out to the central office, program leaders also have built "mid-level" leadership structures to increase the capacity of their programs. Moreover, they recognize the importance of engaging principals and school-level leaders in the science program to provide instructional support to teachers and/or leadership support for the program leader.

Money

- *Supporting a science program with district funds requires vigilance and creativity on the part of program leaders, and commitment from the district's administration.*
- *External funds can boost a program while, at the same time, accentuating existing or establishing new potential inhibitors to that program's sustainability.*
- *Uses of external funds often reflect the interests of the funder and, thus, can influence the shape of the program.*
- *District funds and external sources of support each are associated with particular advantages and challenges that need to be accounted for within the context of the district's culture.*

Many equate program sustainability with a district financial commitment. While there is no question that money is a critical player in a sustained program, its role is far more complex than the simple presence or absence of financial resources. The source of the money, the amount needed, the way it is used at different points in the developing life of the program, and finally, the nature of district culture and interactions with regard to money all are significant issues.

Funding for each of the science programs in this study has been a complex amalgam of resources, including Eisenhower funds, donations from partners, money earmarked for textbooks, external grants, and general district fund line items. Identifying and tracking the varying sources of funds was a challenge, even for some of the program leaders, indicating that the business of securing funds for a program, even when restricted to within-district resources, is a complex job that requires attention and creativity.

There are advantages and obstacles associated with a reliance on internal funds alone as well as with the acquisition of external grants. Regardless of the developmental phase in which a large grant is secured, the influx of money can enable districts to accomplish large tasks in a relatively short amount of time. In addition to the financial benefits of grants, external funds also bring additional independence, stature, and influence to the program leaders. Even as the grants bring opportunities to the programs, the program leaders have to address some challenges associated with the changing ebb and flow of funds. In accessing external funds, leaders have to accommodate funders' guidelines, which may or may not be consistent with their program's needs. Large grants also create the dynamic of "haves" and "have nots" within a district, and the end of those resources can be perceived as being a loss for the program.

Districts in this study that avoided the problems of seeking and receiving external funds have taken pride in their self reliance. Although funds have certainly fluctuated in all of these places, the science programs are accepted practice and, thus, receive consistent support. What leaders gain in avoiding the pitfalls of external funding, however, they lose in the ability to make large-scale impacts on their programs in short periods of time.

These sites suggest that there is no single way or best way to fund a hands-on science program that will ensure its sustainability. Rather, it is the leaders' abilities to understand and address the complex nature of securing financial support that is key.

Partnerships

- *Typical partnerships are somewhat superficial and supplemental but still serve to enrich the science program.*
- *Deep partnerships are rare, require investments of resources and political currency, and can have both positive and negative impacts on the sustained science program.*

Districts in this study have had partnerships that fall into two broad categories. Most common have been the "limited" partnerships forged between a local business or organization and a single school or district area. The other category of partnerships encompasses those that have been deep and comprehensive. Such partnerships are rare, occurring mostly at the district level and requiring investments of resources and political currency, as well as shared planning and leadership. As with many of the other factors found to be significant to sustainability, partnerships are a component that can have positive and negative effects, depending on the context and conditions in a district.

Professional Development

- *The roles of specific approaches to professional development in sustained programs vary, depending on where the programs are in their evolution.*
- *Professional development needs perceived by program leaders are not necessarily congruent with the needs perceived by teachers, nor are they necessarily the activities that will support the sustainability of the program most effectively.*
- *Professional development contributes to sustained programs independent of its impact on classroom practice.*
- *Teachers trained to provide professional development support at either the school or district level often represent unrealized potential.*

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

The role that professional development plays in sustainability is somewhat unexpected due, in large part, to the fact that its intended impact on actual classroom practice is unknown. Still, it appears to have an unintended but no less significant relationship to the sustainability of the programs in this study. This is primarily due to its ability to foster deeper understandings of and commitment to the programs' underlying purpose. This was particularly true for teachers who had anticipated in "higher-level" professional development because they immensely appreciated the messages of respect and professionalism that were implied through their participation in those events.

SECTION 3: FACTORS THAT PERTAIN TO THE WHOLE SCIENCE PROGRAM

Adaptation

- *No district is static. Thus, science programs must adapt if they are to endure.*
- *Sustained programs are altered in a wide variety of ways for a variety of reasons.*
- *Adaptations can be proactive or reactive.*

The definition of sustainability presented in this study suggests that sustained programs use their core beliefs and values to guide adaptations to change. The earlier discussion of what sustainability is and the phases that programs move through asserts that programs must move beyond establishment and maturation of a particular design to a state of evolution in which elements of the program can vary greatly from the program as originally conceived. It is in this movement—from maturation to evolution and beyond—that programs demonstrate the flexibility and resilience essential to their survival in the ever-changing and, sometimes, volatile district environment. Indeed, every program in this study underwent adaptation.

Some of the most visible adaptations are evident in changes to the instructional materials themselves and in their distribution systems. Other less obvious but still concrete adaptations to curricula focus on the instructional sequence. Another common area of program adaptation is the design and focus of program professional development support, which occurs for a range of reasons (including changing district priorities, leaders' changing views of high-quality professional development, and most often, the arrival of external funds). This illustrates the point that adaptations can be proactive or reactive.

Less tangible adaptations also guided the evolution of the sustained programs in this study. Program leaders make adaptations to the program goals, expected outcomes, and their own personal understandings about the extent to which the programs could and should purely reflect inquiry-based instruction. Adaptation in program goals and intent are sometimes subtle and evident only in retrospect, even to the leaders themselves. They sometimes emerged only when looking at a collection of program elements over the long-term time horizon of those places that had operated for 20 years or more. Leaders of younger programs can benefit from the recognition that program goals naturally will evolve and adapt to shifting district conditions and contexts, turnover of leaders, and trends in funding sources. However, recognizing that the core beliefs and values do not waver throughout all of the adaptations is key.

Critical Mass

- *Considering critical mass through the long-term time horizon of sustained programs sheds light on alternative views of what critical mass is and how to achieve it.*
- *In the relative short term, attention to critical mass is highlighted by the challenge of reaching sufficient numbers of teachers.*
- *In the relative long term, attention to critical mass is expanded to include the challenge of obtaining widespread and deep commitment to the core values of the program.*

Discussions of critical mass in reform programs often focus on numbers: numbers of teachers participating, numbers of students reached, and the resource-to-teacher ratio. This is consistent with a view that one 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. The definition of sustainability generated by this study expands this view to suggest that a program reaches critical mass only when there is a culture of program self-generation. Thus, "critical mass" can encompass other considerations more complex than the simple act of targeting a "magic number" of teachers to implement the program. The data of this study suggest that it also is meaningful to consider critical mass as numbers of teachers and principals who understand and believe in the program's core beliefs and values.

This is not to suggest that breadth of training is irrelevant to sustainability when compared with depth of belief in a program's core values and beliefs. Rather, these two aspects of critical mass are intertwined, with one requiring more emphasis than the other, depending on where the program is in development. Clearly, breadth contributes to the culture of program self-generation in an ongoing fashion, particularly in the relatively short-term time horizon. However, when programs have experienced shocks, depth of understanding has played an important part in their sustainability.

Perception

- *The perception of a science program can differ greatly from the actual status of that program in a district. "Misperceptions" can both contribute to or inhibit the sustainability of a program.*
- *In the absence of firsthand knowledge of the status of the program, program leaders and other decision makers take action based on their perceptions.*
- *There is a disconnect in perceptions of the status and importance of the program held by stakeholders at different levels. This confounds efforts to accurately diagnose and address needs.*

Perceptions—whether held by program leaders, program participants, or outsiders to the district—have the potential to significantly support and/or inhibit sustained programs. In some cases, perceptions of the programs differ greatly from the apparent actual status of the program. This is significant because, in the absence of enforced accountability measures, perception becomes 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 strong and exemplary, fostering a kind of complacency. While this impression is positive, it also opens the door for potential neglect in allocations of future district dollars and attention.

Given the lack of authentic data on the status of a program, perceptions of it are often all that decision makers have to guide their actions. The fact that there are disconnects and misperceptions at every turn make the challenging job of growing a districtwide science program even more difficult. It also suggests that perception has been sufficient to sustain these nine programs up until now. In an environment of increased scrutiny, however, it is impossible to say whether perception alone will continue to be adequate.

Philosophy

- *In sustained programs, there is a widespread, shared philosophy that science should be taught using a hands-on approach.*
- *Science programs become vulnerable in the presence of inconsistent philosophies about the importance of teaching science.*
- *The growth of the hands-on philosophy is supported when there are pre-existing or newly emerging complementary approaches elsewhere in the district.*

This study demonstrates that philosophy, a set of beliefs about the role of and appropriate pedagogy for science in elementary education, as it was expressed by teachers, principals, and administrators in the sustained programs, falls into two categories: (1) beliefs about the importance of teaching science, and (2) beliefs about how science should be taught. These two philosophical strands evolve, sometimes together, sometimes independently. In sustained programs, the second strand, relating to how science should be taught, is consistently strong—educators in these districts articulate beliefs that the hands-on approach to science instruction is the best way to teach science. However, the first strand, representing belief in the importance of teaching science at the elementary level, fluctuates depending on the changing district conditions. Thus, even though the programs demonstrate widespread common beliefs about science instruction, they remain vulnerable when lacking support for making science a core part of the elementary student instructional experience.

Although the two strands of philosophy are related, they are not mutually dependent. It was not uncommon to find districts where the commitment to teaching science had varied greatly over time, while the belief in teaching science with kits remained strong. In the face of a focus on other subjects, the science programs in this study sometimes adapted, adjusted, held their ground, or even retreated somewhat, while still holding fast to this second strand, the importance of teaching science using a hands-on approach. The belief in the importance of teaching science must be extremely strong to withstand the pressures that come from accountability for other issues. Generally, that belief has been strong enough to sustain programs through challenges, but not necessarily strong enough to give the sustained programs a sense of security. Only when both strands are strong does program vulnerability fade.

Quality

- *There are no effective mechanisms in place for assessing the quality of science instruction and/or the impact of professional development.*
- *In the absence of accountability measures, actual student learning of science concepts and processes becomes irrelevant to a program's sustainability.*
- *In the presence of accountability measures, program quality is defined by evidence of student performance on those accountability measures. Thus, the degree of alignment between the program and the district's accountability system becomes the primary indicator of program quality.*

This study defines the quality of a program as the extent to which its instruction and curriculum 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. If the quality of curricula and instruction is to have an impact on a program's sustainability, there must be mechanisms in place that allow program leaders and others to gain and maintain an understanding of their status. What is most striking, is that none of the districts in this study have any such systems in place. It is impossible for any of the program leaders to have a sound understanding of the quality of instruction or the impact of professional development on classroom instruction.

Over the course of their programs' history, several leaders have made attempts to understand the status of their programs, and their findings corroborated the findings of this study: Implementation of kits within each district is uneven, and, when teachers do use kits, their practice is highly variable. Leaders are also ill equipped to assess the impact of the professional development they provide.

These findings imply that the quality of instruction and professional development is irrelevant to a program's sustainability. Until the recent past, central office administrators and the general public placed relatively little emphasis on assessing programs' quality, as long as the program was seen to function smoothly with no complaints. Since the importance of student achievement on standardized tests has taken hold, the definition of quality has come to mean student scores on science tests. In this environment, the possibility of bypassing hands-on curriculum in favor of textbooks becomes more attractive to teachers and principals. The implications for future sustainability are worrisome.

CONCLUSION

Throughout this research, program leaders expressed the hope that a consistent pattern would emerge from the data collected across these nine programs and offer a formula for sustainability that would guide their efforts. They dearly wanted more knowledge about how to maintain their programs, strategically concentrate their efforts, and build capacity for continuous growth and improvement. However, as evident in the preceding discussion of the findings, no such formula emerged. Rather, this study identified a set of factors that affect the sustainability of hands-on science programs in fluid and interrelated ways. The roles these factors play in reform efforts are greatly varied and change over time and from place to place as they reflect the complex district environments around them. Within this complexity, while there is no formula for sustainability, the factors presented here illustrate trends that offer new insights into sustainability for program leaders, district administrators, and funders as they invest in new and ongoing reform efforts.

THE FACTORS

The factors that support and/or inhibit sustainability of districtwide hands on science programs fall into three categories: those that pertain to conditions surrounding the district and its program, those that pertain to individual components of the science program, and those that pertain to the program as a whole. See Table 1.

Table 1

Factors that Pertain to Surrounding Conditions	Factors that Pertain to Science Program Components	Factors that Pertain to the Science Program as a Whole
Culture	Accountability	Adaptation
Decision Making and Power	Implementation	Critical Mass
Science for All	Instructional Materials	Perception
	Leadership	Philosophy
	Money	Quality
	Partnerships	
	Professional Development	

IMPLICATIONS

These findings offer many implications for program leaders, district administrators, and funders with regard to their investments in their science programs. Some are described below:

- *Leaders and supporters of districtwide programs can gain from giving attention to the wide range of factors that affect sustainability and account for them in all strategic and financial decisions.*

The leaders of the sustained programs in this study emphasized the factors related to program elements throughout the lives of their programs—even as they moved out of the establishment phase and into maturation and evolution. Thus, their investments in and accounting for the other factors have been a fortunate by-product. This study serves current and future program leaders by making the factors that pertain to surrounding conditions and the whole program more explicit, allowing leaders to be more purposeful about how, and in what ways they allocate their resources.

Thus, this study offers readers an illustration of the importance of attending to factors that are not often addressed or even recognized as important to sustainability of a program. It highlights the concrete ways that the programs in this study have done so, albeit most often unintentionally, and offers a starting point for systematically assessing the importance of each, given the particular time and circumstance, and developing strategies to accommodate them.

- *Leaders and supporters of districtwide programs can benefit from defining and considering sustainability through the lens of a long-term time horizon.*

The RSR project's definition of sustainability, discussed in detail earlier in this report, while acknowledging the factors that pertain to program elements, highlights the important contributions of the factors that pertain to surrounding conditions and the program as a whole. In particular, it refers to the significance of core beliefs and values (philosophy) and adaptation, and acknowledges the importance of culture and decision making and power as sources of change and pressure:

> *Sustainability: The ability of a program to maintain its core beliefs and values and use them to guide program adaptations to changes and pressures over time.*

This definition of sustainability stresses a shift in understanding from sustainability as program maintenance, in which the elements of the program are preserved over time, to one of adaptation, in which the program elements evolve and adjust. Thus, a look at reform through the lens of this definition of sustainability suggests that it is appropriate to reconsider expectations for the outcomes of program investments. Educators need to recognize that change can be subtle and it can be latent. And simply because there is no evidence of a "revolution" does not mean that there isn't important evolution. Educators are well-served to reserve judgment about the failure or success of reform until considering all of the ways it may have affected educational practice and interpreting evidence of those changes in light of a long-term definition of sustainability.

- *Leaders and supporters of districtwide programs must increase attention to the quality of their programs with explicit, focused strategies.*

Hand in hand with discussions about how to sustain programs, educators also should engage in a careful and critical look at what is being sustained. The programs in this study were unevenly implemented and, as such, did not represent the districts' articulated goals for their districtwide programs. Even programs at the height of their renown were not as thoroughly implemented as their reputations would have suggested.

The issue of program quality is of obvious critical importance to all stakeholders but faces obstacles that prevent leaders from both assessing its value as well as improving it over time. These obstacles principally grow out of the lack of leaders' authority, limited access to classrooms, and lack of capacity to collect data and make use of it.

Evidence is essential and beyond the reach of program leaders: Sustaining a program of high quality requires evidence of its impact and its status. The inability of leaders to gather such data is stunning in its absence and chilling in its implications. Lack of evidence of student outcomes, as well as evidence that is not aligned with the goals and intent of the program, leave it vulnerable to being misunderstood and undervalued. In the same vein, without knowing how students are progressing, it is impossible for leaders to know how to direct program improvement efforts.

High teacher and principal turnover locks leaders into the cycle of continuous re-establishment and limits their ability to attend to quality issues in the long-term: Districts that were characterized by a high degree of stability were far more able to advance from the establishment phase into maturation and evolution than districts where teachers and administrators came and went through a revolving door. While they, too, struggled with questions about the quality of that professional development and its impact on classroom instruction and students, they were better equipped to develop strategies for addressing quality concerns.

If educators accept the premise that professional development is linked to quality of instruction and program implementation, they must recognize the challenge of teacher turnover and account for it if sustained programs are to offer high quality instruction that promotes student learning. Regardless of the approach, district and program leaders can not avoid the need to address the threat to stability and lost investments posed by high teacher and principal turnover rates.

IMPLICATIONS FOR THE FUTURE

This study makes the evolutionary nature of reform programs, as well as the patterns of disturbances that they endure, explicit. The shocks and pressures that influence programs' sustainability, such as a change in a district's financial status, a shift in public demand for accountability, or decentralization, are standard fare and, in response, all districts experienced ebbs and flows in the strengths and capacities of their programs over time. Some programs waxed and waned dramatically, but history clearly showed that all programs, regardless of their age or apparent stability, were vulnerable to shocks and pressures, the majority of which were beyond the control of the program leaders. And yet, given society's propensity to debate the value of and need for reform efforts and even specific approaches to instruction, any expectation that a sustained program would become immune from these challenges is misguided. Sustained programs are noteworthy not because they have eliminated threats, but because they survive in the face of them.

This study finds that sustained programs withstand these potential threats with resilience that lay in strengths not easily seen. They were in places where no one had looked—meaning in the more subtle factors of adaptation, perception, philosophy, and critical mass—and were apparent only after the passing of time. Understanding sustainability from the perspective of history and these more subtle factors does not guarantee better outcomes for hands-on programs. But, it does argue that, if leaders attend explicitly to what were previously unrecognized program supports, as debates arise about the way science should be taught and the worth of hands-on programs, their value will be explicitly and thoroughly presented. Likewise, when more hospitable times return, programs will be better equipped to advance further, with greater confidence in their awareness of the gains they have made.

PART I: RESEARCH QUESTIONS AND METHODOLOGY

RESEARCH QUESTIONS

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 comprehensively improve their science education programs. This study addressed 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, as it came to be known, 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 districtwide approaches to science education reform as well as to the role that external funds can play in initiating reforms that are sustained. The research, supported by a grant from the National Science Foundation (NSF), was conducted between 1998 and 2002 by staff of the Center for Science Education (CSE) at Education Development Center, Inc. (EDC), in Newton, Mass., in collaboration with staff of the Caltech Pre-College Science Initiative (CAPSI) in Pasadena, Calif.

The primary purpose of the RSR 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 sustainability and improvement. Specifically, this study focused on nine communities with K–6 science education programs begun from nearly 10 to 30 years ago. This study investigated how and the extent to which these communities have sustained their science education programs and the factors that contributed to their sustainability.

In addition to the global research question: *What contributes to or inhibits the sustainability of a districtwide, hands-on inquiry science program?* the following subquestions also guided the 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?

METHODS OF INQUIRY AND DATA SOURCES

Study Design Overview

The RSR project used a multi-site case study methodology employing primarily qualitative methods, supplemented with a survey analyzed using quantitative techniques. Data collection focused on documenting 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; in-person and telephone interviews of teachers, principals, central office administrators, and other stakeholders; and field notes from observations of classroom instruction and professional development sessions. The data focused on the following areas: program goals and rationale, curriculum, professional development, physical resources, leadership, financial resources and management, accountability and assessment, communication, partnerships, internal and external contexts, and district capacity. This approach supported the team's efforts to gain insight into the political, social, educational, and economic factors that have had an affect on the site and its science program.

Site Selection

The study began with 10 school districts that had districtwide hands-on, kit-based elementary science education programs. The school districts represented a range of sizes, geographic locations, and funding sources as well as varying years of participation in reform. The science kits typically used in these programs included a teacher's guide and accompanying materials for students to use in their science investigations. The kits generally focused on a limited number of core science concepts and contained 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 delivered the kits (usually one or two large plastic tubs) to classrooms and retrieved them at the close of the unit, when they refurbished and prepared them for delivery to the next set of classrooms. In some cases, the kits were housed at the school sites and the materials center delivered replenishments of the consumable items. On average, a class used three to four kits during the school year, and kits rotated from one set of classrooms to the next.

Regarding the date of program origin, the districts fell into two groups: those that began science education reform efforts in the mid-to late 1960s and early 1970s (Bayview, Bolton, Montview, and Hudson) and those whose efforts began in the mid–1980s and early 1990s (Benton, Garden City, Glenwood, Lakeville, Sycamore, and Portman). Regarding funding sources, the districts fell into three groups: those that had no or relatively small NSF funds (Bayview, Garden City, Montview, and Sycamore), those that had NSF Teacher Enhancement funds (Benton, Glenwood, Portman, and Lakeville) and those that had NSF LSC funds that were scheduled to be expended by the time the research ended (Bolton, Glenwood, and Hudson [Glenwood had both TE and LSC funds and received additional funding during the course of the research]). See Table 1 below.

Table 1[1]

School District	# Students K–6 or K–5*	# Schools K–6 or K–5*	Year Program Began	NSF TE Funds	NSF LSC Funds	No NSF Funds
Bayview	5,849	23	1966			x
Benton	4,300	15	1994	x		
Bolton	27,000	60	1977		x	
Garden City	28,000 (K–5)	52 (K–5)	1989			x
Glenwood	27,000 (K–5)	77 (K–8)	1989	x	x	
Hudson	43,151	50	1974		x	
Lakeville	12,000	23	1986	x		
Montview	47,087	92	1968			x
Portman	11,603 (K–5)	28	1988	x		
Sycamore	6,400	30	1988			x

* Figures are approximate for years ranging from 1998–2000

SITE VISITS

Researchers conducted several site visits to each of nine[2] sites (Bayview, Benton, Bolton, Garden City, Glenwood, Lakeville, Hudson, Montview, Sycamore). Teams of two to four researchers typically conducted the site visits with some sites overseen by CSE (Bolton, Garden City), some overseen by CAPSI (Bayview, Glenwood, and Hudson) and others shared

---

[1] All names are pseudonyms.

[2] Portman excluded itself from the study in the middle of the first year due to the arrival of a new superintendent and shifting district interests.

between the research groups (Benton, Lakeville, Montview, Sycamore). RSR researchers organized the site visits in collaboration with site leaders (most often the site's science coordinator.) The research teams provided site leaders with sample site visit schedules that illustrated a model visit for a two-person team. This sample provided guidance for site coordinators and allowed them to devise a schedule that would include key interviews and school visits.

The original data collection design called for each site to have two to three five-day visits. However, based on conversations with site leaders regarding schedules of professional development and other events and on conversations within the research team about the most effective approaches, in some cases, site visits took place more frequently but for shorter periods of time. Site visits in the third year differed from those conducted previously in that they extended data collection beyond gathering information about the program's history and current status. These visits were shaped by the emerging themes identified through ongoing data analysis and included questions designed to enrich researchers' understandings of those themes that were particularly pertinent at each site.

Pre-Visits

The research team conducted a preliminary site visit referred to as a "pre-visit" for each participating site. These visits took place in person, or by telephone when in-person discussions were not feasible. The research teams structured the pre-visits to establish strong working relationships with the site leaders, obtain an initial history and timeline of the evolution of the science education reform effort, identify critical events and major players in the program, and work with the site contacts to identify the most effective use of site visit time.

Interviews and Focus Groups

During site visits, researchers conducted in-person interviews and focus groups to gain an in-depth understanding of individual and group perspectives on issues pertaining to the science program. Data collection typically included interviews with the superintendent, assistant superintendent for curriculum and instruction, director of professional development, director of assessment and evaluation, science coordinator, director of the materials center, resource teachers, board members, union representatives, significant partners or collaborators, classroom teachers, principals, and other appropriate central office or school personnel. In some cases, researchers conducted telephone interviews with individuals who were not available at the time of the site visit. For a summary of the interview data collected across all sites, see Appendix A.

Interview Protocol Development

The research teams began the process of developing the interview and observation protocols by examining other protocols used for similar purposes and discussing early drafts with fellow researchers. Then, researchers tested the protocols through telephone interviews with individuals who have served as superintendents, principals, teachers, and budget managers, and then revised and adapted them as necessary. The protocols were modified after completing the first round of site visits. Interview protocols were designed to gain information about the goals and vision of the district science program, the actual classroom practice, professional development, support for teaching science, and other key critical issues that had an apparent impact on the science program or the district. Interview questions varied according to the individual or group being interviewed. The interviews also explored the factors that interviewees thought contributed to the sustainability of the science program, the factors they thought would support or jeopardize the program, and what they envisioned for the future of the science program in the district. Individuals also were given the opportunity to discuss any other issues they thought were relevant that the interview had not explored. Sample interview protocols are included in Appendix B.

Classroom Observations

The site visits included classroom observations focused on helping researchers better understand the status of the current science program in the district. The objective of the observations was to obtain a "snapshot" of instruction, to identify the school district's practices and goals, and to document the use of hands-on investigation and/or inquiry methods of teaching science. Researchers typically observed entire science classes in grades K–6 that varied in length from approximately 30 minutes to an hour.

The classroom observation protocol was developed based on researchers' previous experience with classroom observations and through making observations during the initial site visits. The protocol was used to guide researchers in making observations in a consistent and systematic manner and examined the following: demographics of the classroom, student work setting, classroom appearance and layout, materials used by students and teachers, the structure and format of the lesson (introduction, body, wrap-up), how students spent class time (e.g., minutes devoted to specific activities), interaction among students, interactions between teachers and students, student engagement, student writing, and assessments.

Researchers typically observed 10 classes distributed across at least five schools during each visit for a total of approximately 20 lessons observed in each site. Schools were selected to represent a geographic and demographic range. Initially, researchers asked to see classrooms that represented "best practice" in the eyes of the program leaders in order to gain a clearer understanding of the district's goals for instruction. After the first round of visits, researchers asked to observe classrooms that represented the kind of instruction that program leaders hoped to achieve and sustain districtwide, rather than the ideal. Some of the observations were made in the presence of the science coordinator, resource teacher, or other district person, and in these cases site visitors discussed the lesson with them after it was complete. These discussions provided additional data relative to the district's perspective on science instruction.

Document Review

The document review included an analysis of all materials that a site coordinator and researchers deemed relevant to the science program in the district. A document and data checklist was developed as a tool for researchers to use with district coordinators in order to prioritize information that they could obtain about a district before making a site visit. The checklist consisted of a list of sources of key information, including budget information, proposals, strategic plans, enrollment records, individual school information, organizational charts, curriculum documents, assessments, testing standards, tests, and promotional literature about the district and the science program. Site contact people used the checklist to familiarize themselves with the type of information researchers would require for the study. The document review was an ongoing process and researchers continued to collect relevant materials from sites throughout the project. The document and data checklist can be found in Appendix C.

Other

Researchers also observed professional development activities conducted by the schools and the districts. These ranged from introductory kit trainings to on-going, in-depth sessions on inquiry instruction. Researchers took extensive field notes, wrote memos on these observations and incorporated this information into the continuing data analysis.

DATA ANALYSIS

Qualitative data analysis was ongoing and iterative, and took place in several stages that corresponded roughly with the site visits. The first stage involved assembling and absorbing data among the members of each site's research team, which consisted of gathering and preparing field notes, sharing observations, and reflecting on preliminary findings. Most on-site interviews were audiotaped, and selected interviews were transcribed. This stage enabled researchers to identify additional people for follow-up telephone interviews and to gather additional documentation from the site contact person that would aid in writing the report. The follow-up interviews provided clarification of earlier identified issues as well as insight into new questions generated from the discussions.

The second stage involved sharing site-based data among CAPSI and CSE researchers through written reports and participation in whole team meetings. For each site, one team member drafted an initial report with input and feedback from the other research team member(s) who participated in data collection at that site. These initial reports were written for the purpose of sharing data and were structured according to an outline that all researchers felt would usefully fit and organize the data. The categories included context, history, status of the program, status of instruction, status of professional development, communication, equity, community and partnerships, articulation, funding, leadership, accountability, capacity, and overall impact on sustainability. Researchers also wrote questions in every section to guide future data collection via site visits, additional follow-up interviews, or document analysis. Reports tended to vary in emphasis, reflecting the varied sites and corresponding data.

Researchers from CAPSI and CSE also met together to discuss their findings, share reports, and discuss strategies for the next round of site visits. At this time, researchers reviewed any problems that occurred or improvements that needed to be made in the research methodology. Most important, researchers discussed individual sites and identified specific areas that required clarification and further data. These reviews formed the guidelines for the final round of data collection. As each new wave of data was collected, the site reports were revised to include the new information. The final site reports are available by contacting either CSE (for online access, visit the CSE Web site at http://www.edc.org/cse) or CAPSI. Executive summaries of each report can be found in Appendix D.

The third and final stage focused on conducting a full analysis of the data across all sites and, in the process, conceptualizing this cross-site report. To accomplish this, researchers examined the original research questions across sites to capture themes and issues that emerged, discussed overall themes and perspectives, and challenged early ideas. The overall objective was to generate a cross-site analysis that would reveal new insights about sustainability.

SURVEY

The RSR project designed two surveys—one for principals, one for teachers—to supplement the qualitative findings of the RSR study by providing additional data on the current status of the program. The surveys were shaped by the following areas of interest: (1) What are the respondents' understandings of the current science program? (2) What importance do respondents place on the science program, and what priority does it get? (3) What are the respondents doing to implement or support the science program? (4) What factors are important in sustaining an effective science program? Specific items focused on teacher and principal background and experience, principal and teacher administrative and instructional practices, curriculum and materials, professional development, influences on science, support for science, and sustainability of the science program.

Survey Design

Survey development followed a three-step process. As a first step, researchers investigated the literature on principal and teacher surveys, including the Third International Mathematics and Science Study; Center for Learning Technologies in Urban Schools (School of Education and Social Policy at Northwestern University); LAAMP School Family Research at UCLA; and Johns Hopkins University (Education in the Middle Grades: A National Survey of Practices and Trends). This review of instruments guided researchers in the type of items and topic areas they should include, the level of information they should seek, and the type of scales they should use in their own survey development.

Next, several revised drafts of the principal and teacher surveys were produced and piloted to establish the validity of the instruments. Based on the four areas of interest outlined above, researchers designed the surveys so that most questions allowed for the comparison of principal and teacher responses within a district. Researchers piloted the surveys with samples of teachers and principals who were not part of the research study, but who were selected only to complete the questionnaires and to provide their feedback and suggestions for improvement. After the questionnaires were piloted, researchers conducted both in-person and telephone interviews with the pilot-test teachers and principals to gather their perspectives on the survey content. This feedback focused on length of the questionnaires and amount of time required to complete them, clarity of the questions, the relevance of the items to teachers and principals, their typical responses, discussion of open-ended responses, and general suggestions for improvement. As a final step, a university researcher and survey expert from the University of Southern California reviewed both surveys. The survey instruments are included in Appendix E.

Survey Administration

The surveys were administered to all principals in eight of the district sites (from 14–77) and 100 randomly selected elementary teachers in each of those same eight sites (6–50 percent of the total teacher population, depending on the site). The response rates for the principal survey ranged from 49–75 percent. The response rates for the teacher survey ranged from 40–82 percent. One site (Montview) declined participation in the survey process due to complex internal approval procedures and other research taking place at the same time. Site coordinators produced lists of principals and randomly selected teachers. The purpose of the survey was to supplement the qualitative findings of the RSR study by providing additional data on the current status of the program. Depending on the response rate and the reliability of self-reporting, these data may not accurately reflect actual districtwide practice.

To initiate the survey process, a memo was sent to the science coordinators to discuss the survey plan and to gain the necessary permission from the district. If a site visit to a district coincided with the survey being administered, researchers took this opportunity to discuss the surveys in person with the science coordinators and gain their input in this process. Teachers and principals received their surveys the first week of January 2000 with a requested February 1 return date. Surveys were accompanied by a cover letter and endorsement of either the district science coordinator or the superintendent. To provide confidentiality, surveys were given an identification number and distributed with self-addressed stamped envelopes so that participants could mail them back to CSE or CAPSI individually.

Survey Analysis

A total of 1,029 surveys were mailed: 800 teacher surveys and 329 principal surveys. After the return deadline, those teachers and principals who had not yet responded to the survey were sent a second survey and cover letter. A total of 683 surveys were received: 213 principal surveys and 470 teacher surveys, returning a response rate of 66.4 percent overall—64.7 percent return rate for principals and 58.8 percent return rate for teachers.

CAPSI and CSE collected the surveys, which CAPSI then coded and entered into an SPSS data file. They also coded the open-ended responses. Frequencies were calculated for each question across all principals and teachers, and again by all principals and teachers within each site. Principal and teacher data were compiled by site, and reports were generated that enabled comparisons to be made between principal and teacher responses to common questions. These reports are included in Appendix F.

PART II: SETTING THE STAGE FOR FINDINGS

THE STORY OF A PIONEER

The site reports of this study describe nine communities and their science programs in detail. We recommend that the reader review these reports in order to fully understand the presentation of the findings in Part III. With this said, we understand that it also will be helpful to have an easily accessible reference for interpreting the findings of the report as they are presented. Thus, we have chosen to include a condensed version of one of the site reports here (other executive summaries of the site reports can be found in Appendix D). We chose Bolton because, with its program in place for more than 30 years, it is one of the "pioneers" of districtwide, hands-on elementary science, having experienced times of strong growth as well as periods of decline and stress. The Bolton story is one of strong, creative, and strategic leadership; development, improvement, and management of instructional materials; and thoughtful, strategic acquisition and management of resources. Its richness in strategies, creativity, and flexibility are overshadowed only by the striking, sound, and steady commitment all of the Bolton leaders have made to hands-on elementary science instruction. This abbreviated story will set the stage for the reader and provide a practical reference for grounding and interpreting the findings described below.

THE BOLTON STORY

The Bolton School District's (BSD)[3] hands-on, kit-based elementary science program was a pioneer in the field and has been a key feature of the academic program since the mid-1970s. At that time, Pearl North, a high school science teacher, introduced hands-on science to BSD and, over the course of several years, established a districtwide program. By 1996 the second generation of the program began, led by Dorothy Parson, a former "teacher expert" who reinvigorated it by redesigning the curriculum and expanding the program's depth and breadth. Unlike North, Parson had the assistance of a core team of teachers and a $3.1 million LSC grant from the National Science Foundation (NSF) that targeted districtwide professional development for elementary science. The third and newest generation of the program currently is led by two teacher experts, Sophia Harder and Maria Clay, who took over in 2000. Both Harder and Clay were key players in the NSF grant and bring years of elementary classroom experience and knowledge of science education to the task of moving Bolton's program into the future. Their story and the story of BSD offer a lesson in program evolution and how each stage of historical development contributes to long-term program sustainability.

Community Overview

Bolton, with a population of 225,000 is among the largest cities in its largely rural state and serves more than 40 percent of the state's children. In 1999, 60 elementary schools in Bolton served over 27,000 students, with 1,300 certified staff and 1,144 classroom teachers. There are 11 middle schools, 6 high

---

[3] Any individual, organization, or corporation named in this report has been given a pseudonym.

schools, and 8 special program schools. Although the district encompasses the municipality of Bolton, which is more than 1,000 square miles, the primary populated area covers about 320 square miles. The remaining area is rural and sparsely populated.

The student population of BSD is moderately diverse. In the elementary grades alone (K–6) (as of 1999–2000), white students compose 62 percent of the population, Native American students 13 percent, African American and Asian Pacific Islander each 9 percent, with Hispanic children accounting for 6 percent. In addition, nearly 11 percent of elementary students are learning English as their second language. Slightly more than 30 percent of BSD elementary students are eligible for free and reduced-price lunch and several thousand students have special needs. Average student mobility over the past several years has remained steady at about 20 percent per year.

Program History and Development

Program Origins: In 1974, Pearl North came to BSD as the second-ever science coordinator. At that time, the elementary science program was a "hodgepodge," and North had a mandate from the former director of elementary education to create "one science program," and from the principals to "do something with all the science junk in the back closets!" North set about this task by educating herself about elementary science and attending many NSF-funded institutes that provided information about different approaches and resources. She eventually learned about a program in a nearby state and went there to talk to its director to learn more. "That was the turning point for Bolton," she recalls. Over the next several years, North modeled her work after the district she had visited and enlisted its director to visit Bolton and advise her and a committee of principals she had organized.

In 1975–76, using commercial kits as a model, North set out to develop kits tailored specifically to the BSD schools. She went to the elementary buildings and emptied their closets of old science resources and engaged BSD teachers and interested organizations of the community in the process of designing and creating the kits. North recalls this time as a whirlwind of activity all focused on getting kids' hands on the "stuff of science." In December 1977, she delivered the first batch of kits to seven volunteer schools, each of which had committed to providing financial support for a shared "resource teacher" and an aide. Over the next two years, the program began to solidify. Use of kits expanded to 38 schools by the end of 1979, and by the very early 1980s, all elementary schools were using the kits on a voluntary basis. Then in 1981, the program passed a key landmark. The Elementary Science Curriculum Committee recommended to the school board that they eliminate the use of elementary science textbooks for the science materials adoption and use kits exclusively. The school board agreed, and with that decision, the science kits became the official elementary science program for the BSD.

In contrast to the previous upward trajectory, the mid-1980s brought a decline in the program. Between 1986 and 1988, the superintendent, who was new to the district, cut many millions of dollars from the budget, and the impact on the elementary science program was profound. Staff and refurbishment supplies were cut, and by 1985 the eight resource teachers that had been in place were reduced to four. In 1985–86 there were three; in 1987–88 there were two; and then in 1988–89, the resource teachers were eliminated altogether. Still, in 1987, even as budget cuts were being made, the science program went through a revision. As part of the district's regular cycle of curriculum review, the remaining resource teachers initiated what was to be a two-year process to improve the quality of the kits. They established a template for each lesson, "tweaking some and substantially expanding others."

By fall 1989, North and the materials center staff had managed the program on their own for a year—with no direct support for teachers—and the central office administrators recognized that it was time to attend to the diminished science program. Dorothy Parson, the program's future leader, was hired as the district's "teacher expert." Also at that time, recognizing that it was due for an adoption of elementary science materials, the district allocated the elementary science adoption money to refurbish the newly revised kits and increase their number. A program revival had begun.

In 1991, Parson took a team of educators to the NEXT STEPS Conference in Washington, D.C.[4] and, as a result, formed a core team of 24 teachers from across all grades. They began planning the curriculum and, as the work progressed, Parson and her team revisited the kits they had in place. Based on the developing new framework, they revised the kits and added new ones, including some that were now commercially available. Then, in 1992, a grant from the U.S. Department of Education enabled the core team to complete the grade-level framework and thoroughly field-test the new kits.

In an attempt to obtain the funds necessary to fully implement the new curriculum as planned, Parson applied to NSF for a Local Systemic Change (LSC) grant. The proposal was successful, and when the field-testing was completed in 1995, $3.1 million in LSC funds enabled a massive, four-year training effort that would involve every elementary teacher each year. Full scale, mandatory training on the kits began in 1995–1996. Teachers also got support for monthly grade-level meetings.

In 2000, as LSC funds wound down, Parson prepared to retire and, thus, leave her position as leader of that project. The two strong "teacher experts," Harder and Clay, stepped forward to take the reins. Their principal concern rested in making the transition from a large, time-limited, externally-funded project to an internally supported, institutionalized district science program. They, as well as teachers, principals, and administrators, were anxious about the void that would be left by the end of the LSC funding. Before Parson's departure, she reflected on sustainability, the LSC, and what she might have done differently. She observed that although she might handle the next phase of the program differently than Clay and Harder, the program was in good hands and it would (and should) develop in a way that reflected their unique styles and interests rather than her own.

The Current Program

The Global Community Science Program (GCSP), as it has become known, is the district's kit-based, elementary science program. Each classroom teacher is expected to use three kits per year with each kit covering a seven- to eight-week period. The kits are prescribed by the program and cover three strands: life science, physical science, and earth science. A fourth strand, known as *Explorations in Science*, is an opportunity for teachers to explore topics that respond to the particular interests of their classes and/or to relevant community issues.

The science materials center enjoys strong support from the district. It is run out of the district's warehouse and, now, is quickly outgrowing the space. BSD has recently adopted a new computer-based inventory system, revamped the way vendor information and ordering procedures are managed, and developed a sophisticated ordering and distribution system to process teachers' spring kit rotation requests.

In general, teachers express a desire for students to develop their natural curiosity and enthusiasm for doing science. One teacher commented, "I want them to experience science, and these kits allow them to do that," while another said she wanted "to encourage each child to participate and get their hands in it at this age." Another typical teacher comment—"I want the kids to pose their own questions for their own experiments and have the opportunity to test some things"—reflected a general interest in having their students understand the scientific process as well as particular science concepts. Finally, many teachers explained that they wanted to foster the joy of learning and exploration.

---

[4] NEXT STEPS was originally sponsored by the Association of Science Materials Centers and now is jointly run by ASMC and the National Science Resources Center (NSRC).

At the same time, teachers pointed out the challenges of teaching science using a hands-on approach. They referred to the time and effort involved in preparing lessons, and the need to trust students' ability to learn. Listening and facilitation skills also are critical, while increasingly large class sizes and the district's focus on reading and math make it a challenge to find the time to teach the kits. Many teachers complete the three kits that are prescribed by the program, while the fourth strand, *Explorations in Science*, often goes unattended. One teacher captured the sentiment of many when she said, "Teaching reading is "easy" compared with inquiry science."

District Decision Making and Leadership

Decisions about the district's curricula are a core responsibility of the district's curriculum department. The curriculum review process, which is intended to take place on a 7- to 10-year cycle, is quite laborious and inclusive, calling for the involvement of an extensive curriculum review committee. This review committee conducts a detailed assessment of proposed curricula and makes recommendations to the school board for approval or alterations. The school board can accept the recommendations of the committee or act independently, and has done both in the past.

BSD superintendents also have a history of influencing the development, growth, and evolution of new initiatives. For North, their active support paved the way and provided her with the resources she needed to grow the program. Parson was not so fortunate. During her years, the superintendent was not actively resistant, but neither was he overtly supportive. As a result, Parson sought and found her allies in other places and had to make do with meager access to those with ultimate decision-making power. As Harder and Clay continue to lead the program, the superintendent's role remains to be seen.

The BSD principals have long felt limited in their power to make decisions about their school's programming, budget, and resources. They find this quite frustrating, particularly because they are under immense pressure to improve student achievement. While principals have control over professional development programs to improve staff skills, they are hampered by the continual reductions in professional development time. The emphasis on student achievement in reading and mathematics also exerts significant pressure on principals to push teachers in these areas and focus less on science instruction.

Establishing, improving, and providing continuing support for GCSP over the years has required a range of leadership skills each used at the most appropriate time. The first two generations of GCSP leaders had very different styles, but their approaches meshed well with the program's needs and the district context of the time. Looking to the future, Harder and Clay have a strong working relationship and a shared vision of how GCSP should grow. In general, their view of the program's next phase includes deepening teachers' understanding of inquiry, integrating science with the rest of the elementary program, and increasing the likelihood that science will continue to be taught.

Resources and Support

Funding: A key element of a program's sustainability is the extent to which the district steps in to assume program costs that had previously been supported by outside funds. In 1998–1999, GCSP funding included a mix of district and federal Eisenhower funds, as well as remaining LSC funds. Since then, the district has continued to support the science materials center and has used its Eisenhower funds to support one teacher expert (Harder) while district funds support the other (Clay). Many administrators expressed their pride in the district's deep level of commitment to the elementary science program. The assistant superintendent for instruction, for example, explained that they have made a very strong commitment to the science program, particularly in light of the BSD budget cuts that rarely allowed anything more than "maintenance"-level support.

As the LSC project has shown, outside funding can contribute greatly to program development and ultimate sustainability. However, seeking external grants is not without its costs and challenges. In BSD, anyone is free to seek external funds. The district's only grantwriter focuses her efforts on large federal grants. There is no formal process for making decisions about which grants to pursue based on alignment with district goals. Rather, the grantwriter explains, "I look for the opportunities within the district, and I match them with the external opportunities." In many ways, this approach has served the district well. With help from various program leaders with initiative, drive, and skill, she has raised about $40 million of much needed support over the past 12 years.

Community and Partnerships: In the early years, the local community played a significant role in establishing and developing the elementary science program. Partnerships with local environmental organizations were key assets during North's tenure, offering content expertise as well as moral support. These connections to the community have remained, although they are not as strong today. Several local environmental and science organizations have developed program offerings that correspond to the science kits, and some provide space for training sessions and offer a range of experiential programs to the schools in the district. Finally, many of BSD's parents are involved in science or environmental-related work or recreation and, as a result, they seem to appreciate the science curriculum.

Accountability

Accountability for Student Learning: The State Board of Education recently mandated that each local district adopt the state's content and performance standards in reading, writing, and math. It also mandated that all children should be independent readers by the third grade. In response, the BSD approved its own, more rigorous standards. Science content standards had also been approved, but as of April 2001, the science performance standards were "caught in a political mess." If tests are developed for science and/or social studies, the director of assessment predicts that they will be implemented in the eighth grade.

BSD students in grades K–3 receive grades only on "effort" in science, while students in grades 4–6 receive marks on "performance" and "effort." In addition, teachers use assessments that the GCSP developed specifically to target students' conceptual understanding. However, BSD's capacity to pursue its own curriculum-driven assessment has been curtailed as a result of the state's assessment program. Since the state has increased its role in testing, BSD's Department of Assessment and Evaluation has reduced its size and scope. Over the course of the 1999–2000 and 2000–2001 school years, its budget was cut by 45 percent with a corresponding cut in staff. There were no district-level tests in place as of 2000–01. Moreover, if curriculum coordinators want to improve their program's assessment tools, they would have to contract with someone outside the district as the department can no longer provide that expertise internally.

Accountability for Teaching the Program: The commitment of teachers and principals to teaching science is extremely variable. Although the extensive training provided via LSC funds went a long way to address teacher reluctance, it still remains a problem. Furthermore, although it is common knowledge that there are resistant teachers and principals across the district, program leaders are unable to discern the magnitude of the problem. This lack of awareness is due to several factors. First, the GCSP does not have the resources to maintain firsthand knowledge of the quantity or quality of science instruction that takes place in schools. The two teacher experts, Harder and Clay, cannot visit enough classrooms in the 60 elementary schools to say with any confidence that they have an accurate account of the status of the program as it is being delivered. And, although the clerks at the science materials center see for themselves how thoroughly a kit has been used when it is returned, this information is not captured and used at this point. The new program leaders are aware of this information gap and are interested in addressing it in the coming years.

Another stumbling block to ensuring science instruction is the voluntary nature of training for new teachers. Central office administrators were pleased that the district chose to fund the teacher expert positions (Harder and Clay) that had previously been supported by LSC funds, and felt that this was a strong step toward ensuring that training for new teachers would continue. However, there still is no way to guarantee that all the teachers who need training on a kit actually attend the training sessions. Many teachers believed that the only guarantee that all students will receive science is the presence of a standardized test in science that is of equal importance to the tests in reading and mathematics.

Equal Access to Science

Teachers motivated to do so can avoid teaching the kits, and when this happens, the children in those classrooms simply do not receive science instruction. Additionally, when teachers do teach science, not all students are equally engaged. This latter problem explains, in large part, the interest that GCSP leaders have in a program known as the *Kagan Cooperative Learning Project*. This program has varied learning structures that are designed to ensure that all students in a classroom participate, while at the same time, manages the nature of participation so that it is organized and controlled. In general, teachers involved in the *Kagan Program* report a greater sense of control as well as a greater degree of student engagement that can translate to more widespread, authentic involvement with the science lessons.

Summary

Bolton School District's current program is the natural and obvious descendent of the previous generations of elementary science. The outstanding trait that they all share is, perhaps, not really a feature of the programs themselves, but of the school district and larger Bolton community in which they are embedded. That trait is an abiding commitment to having children use materials to study science, and it is striking to note how firmly the district identifies its elementary science program with the use of science kits and the science materials center. The long history of the program shows that there have been periods when the program waxed and waned. The district's strong commitment to kits has helped the program take advantage of opportunities to expand, but it has not protected the program from feeling the effects of shocks, be they due to the budget , teacher turnover, or a focus on other subject areas. BSD will never be immune to unpredicted, dramatic, far-reaching events, but its past experience with survival during turbulent times is instructive. Only with hindsight over such a long period of time can one understand the unpredicted and subtle aspects of the program's sustainability.

MAPPING THE FINDINGS

The story of Bolton gives the reader a brief glimpse into the real world of hands-on elementary science programs and how they endure. However, a full discussion of sustainability as it relates to the findings of this study is more complex and covers a spectrum of factors ranging from those that are explicit and concrete to those that are more subtle. Figure 1 on page 17, Sustainability Factors and Surrounding Conditions, maps the factors and conditions we find to be important to sustainability and represents the complexity of their interactions with one another.

In the center are the most concrete factors that contribute to or inhibit sustainability—those that pertain to specific science program components (accountability, implementation, instructional materials, leadership, money, partnerships, and professional development) and those that are somewhat less tangible and pertain to the whole science program (critical mass, adaptation, perception, philosophy, and quality). The next ring represents factors that influence the conditions that have some bearing on the operation of the program in the district context and on the strategies program leaders employ to support the program's stability and development. These factors are the core findings of this study and are discussed in detail in Part III of this report.

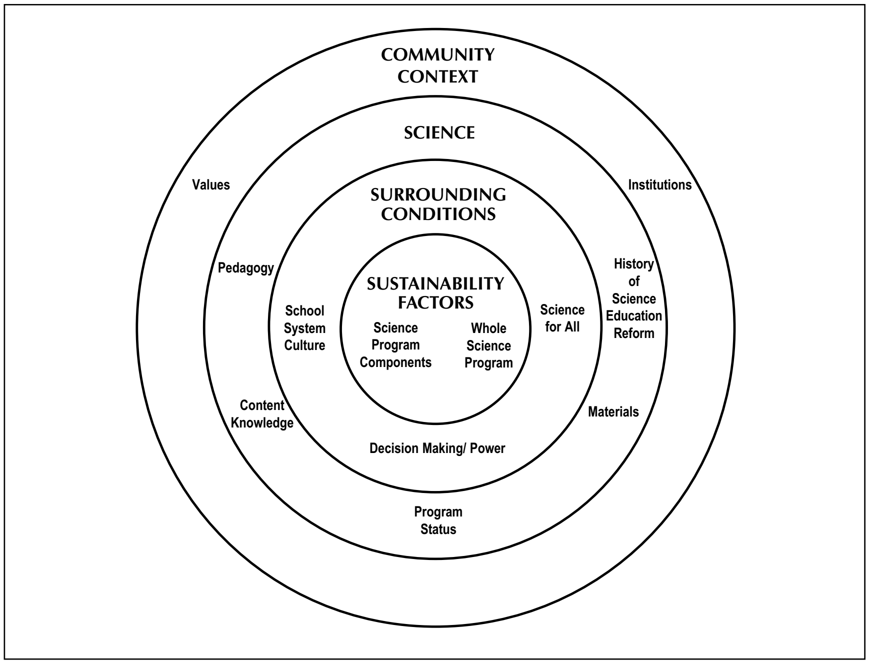


Figure 1
Sustainability Factors and Surrounding Conditions

In addition to the factors that contribute to or inhibit sustainability, we also found that other aspects of the context are likely to play a role in the evolution and endurance of the program. As indicated by the next ring of the map, one such contextual issue is the fact that the programs in this study focused on the subject of science. Science as a subject area, particularly when implemented at the elementary level, brings with it an array of unique issues for consideration.

For example, many elementary teachers enter the classroom with little or even no science background outside of a brief introduction to methods of instruction for science. As a result, they sometimes feel unqualified, fearful, and intimidated by the prospect of teaching science. Their perspectives on science also are shaped by the underlying pedagogy of hands-on science instruction. The programs of the districts in this study espoused a philosophy that supported a high degree of student engagement and self-direction in the learning process. Even when a teacher has a strong content background, this approach to instruction (most often quite different from the teachers' experiences of learning science out of a conventional textbook) takes a great deal of practice and professional development to implement effectively and appropriately.

Instructional issues in science also are complicated by the fact that hands-on instruction calls for the use of materials. A district or school must find a way to identify, purchase, store, and distribute materials while, at the same time, providing training on materials use for teachers. In some communities, just the materials management component alone of a hands-on science program requires a level of effort equivalent to complete projects in other subject areas.

Further, the historical, social, and political context of the district and nation as they relate to science also influence the sustainability of the program. For example, several sites, such as Bayview and Montview, were initiated as a result of or bolstered by the launching of Sputnik, the related national concern about the quality and quantity of scientists in the United States, and the subsequent national curriculum development efforts of the 1960s and 1970s. Others, such as Sycamore and Benton, had their roots in this era but weren't established as districtwide programs until the late 1980s and 1990s within the national environment that emphasized consistent, high standards. Similarly, today's political environment of increased accountability for reading and mathematics affects efforts to establish and develop science programs because, in comparison with the other subjects, the status of the science program is diminished, leaving it vulnerable to reduction of time devoted to it or even complete neglect.

Finally, the broadest ring in the diagram represents the community context. This refers to the values and institutions that are predominant enough in the community to influence the program and shape program leaders' decisions. For example, community members in Bolton communicated a widespread interest in environmental issues that was evident in parts of the program implementation, the early curriculum units, and the supplemental materials. In contrast, Garden City is a community surrounded by a fast-growing population of industry and business. Thus, the Garden City science program has received support from several corporations and continues to seek ways to partner with some of the largest corporations in the area. Although the community context is perhaps the most removed from an elementary science program's daily work, it clearly exerts pressures that can play a powerful role in a program's constancy and growth.

DEFINING SUSTAINABILITY AND THE PHASES OF CHANGE

As the discussion of Bolton, and the map of the study's findings above illustrated, the story of sustainability is complex with many components that act alone and in concert with each other. The broad impact of the community as well as the unique nature of science already has been described and other influential conditions as well as the factors themselves will be discussed in detail below. However, in order for these findings to be coherently understood, they must be clarified by two important concepts: what we mean by sustainability and the way in which science programs experience changes over time.

Defining Sustainability—Maintenance vs. Sustainability

Early on, we recognized that educators commonly viewed program sustainability as program maintenance—embedding a program, as designed, into a standing operating system. This approach suggests, then, that a program is sustained only if it looks and functions the same way 2 years, 5 years, or 10 years after it was first established. By this definition, anything short of a replica is not sustainability.

As we explored the meaning of sustainability in the context of these nine school districts, however, we found that this view of sustainability was limiting. We acknowledged that "sustaining districtwide education reform" was 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 we became familiar with each program, particularly the "middle-aged" and "older" programs, it became clear that none of them were exact replicas of their earliest years. Each program had evolved over time, and the longer the time horizon, the more clearly the trends in evolution emerged from the daily, monthly, and annual activity that all programs experience. Bayview provides an example. There, a natural drive to improve the program and its curriculum was a catalyst for many adaptations over the program's long history. The leaders placed an emphasis on the continual evaluation and revision of the program's kits, which eventually led to a far more structured program than the original. The kits themselves were different and more focused, teachers' use of them shifted from individual choice to a prescribed scope and sequence, the professional development shifted to emphasize different aspects of instruction, and the kit maintenance and delivery system was adjusted to the altered curriculum. And yet, the commitment to the use of hands-on materials never wavered. In fact, the commitment to the highest quality of materials was the very reason the program had changed at all. Opinions may differ with regard to whether or not Bayview was the same program 30 years after it had begun, a near or distant relative, or a completely different entity. But we would assert that the vision has remained constant across leaders and over time, and, indeed, this program is sustained.

Sustainability: The ability of a program to maintain its core beliefs and values and use them to guide program adaptations to changes and pressures over time.

Thus, we have made a clear distinction between program maintenance and sustainability. A program can be considered *maintained* if its basic elements (e.g., instructional materials, professional development program, leadership plan) are well established and are commonly accepted as standard practice. Our definition of sustainability, on the other hand, stresses the importance of adapting and improving in response to the changes that inevitably occur in a school district. 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. But adaptability alone is not enough; adaptations must be guided by the essential values and beliefs that characterize the core of the program's intent. It is the continued influence of those beliefs and values that ensures that, as programs evolve, they remain closely connected to their earlier generations.

Adaptability is essential to sustainability for three reasons. First, sweeping changes occur regularly within districts and exert pressures on educational programs. These changes may offer opportunities, or they may pose obstacles for program and district leaders, but either way, leaders must address them. Examples of such changes, or what some refer to as "critical events" or "shocks to the system," 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. All of the programs we studied experienced events such as these: Just to name a few, Lakeville underwent a shift in political agendas; Bolton experienced a large turnover in teaching staff; and Garden City survived a curriculum adoption. Depending on the particular district and science program, any of these events could represent a serious threat to the science program, an opportunity for improvement, or simply a period of turbulence that must be weathered. In any case, programs that endure must be able to navigate occurrences such as these with flexibility moderated by a continuing commitment to program goals.

Second, there are myriad smaller but more frequent, critical challenges that arise in districts, calling for leaders' attention, creativity, and flexibility. As discussed above, elementary teachers can be reluctant to teach science because they lack familiarity with the science content. Or, even if the content is not a problem, they may be overwhelmed by the need to manage the range of materials and activities in a science lesson. At other times, public attention may turn to student achievement in reading and mathematics, compelling elementary teachers to reduce or even eliminate time for science instruction. And still at other times, district policies, such as the practice of providing minimal to no accountability for teaching science, may enable teachers to avoid teaching a science program altogether.

Finally, as knowledge of teaching and learning continues to grow, so do program leaders' goals and understandings of appropriate expectations for program delivery and student outcomes. Thus, leaders continuously work to improve their science programs and look for ways to incorporate their new understandings. Even though adaptations of this nature are internally motivated and within program leaders' control (unlike the external pressures described above), they are every bit as complex. As the brief discussion of Bayview suggested, each improvement brings with it a ripple effect of changes that must be addressed if an overall advancement is to be realized.

Three Phases—Moving Toward Sustainability Over Time

We identified three stages of program development that advance programs from maintenance to sustainability. While arrived at independently, our stages are compatible with those developed by others who have studied change. We chose to label the stages *establishment*, *maturation*, and *evolution*. Though the programs we studied approached these phases differently, as the site reports demonstrate, each phase was still characterized by some common features. Moreover, the lines of demarcation between phases are not exact; they blur, and programs do not always move forward smoothly. They may advance, hold, slide back, retrench, and then move ahead again. But, as mentioned earlier, the longer a program's time horizon, the more clear its pattern of growth and development.

The *establishment* phase focuses on introducing the use of kits for the instruction of science, distributing the kits to 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 districtwide. 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 must establish lines of communication and support systems so that teachers will be able to report their concerns and leaders can understand and respond to problems as they arise. All of the features of the establishment phase correspond closely to the factor of implementation, covered in detail in Part Three.

Lakeville and Benton are good examples of how districts can approach establishment in very different ways, albeit with similar goals. In Lakeville, one school was selected as a "pilot school" where, in the first year, interested teachers were introduced to the use of science kits, received training on them, and, ultimately, began to teach science with them. Over the course of the next two years, the program grew incrementally. The number of schools and teachers expanded as the infrastructure for handling the materials, professional development strategies and resources, and the program's leadership was solidified. After the pilot period was officially over, the establishment phase continued until it eventually encompassed all of the elementary schools in the district. Although this process was aided by the acquisition of NSF funds, the essential model remained consistent with the incremental pilot-school approach.

In contrast to Lakeville's strategy of starting small and growing slowly over time, Benton began its establishment phase with a large NSF grant. These funds allowed the science program to commence districtwide by hiring five district resource teachers to simultaneously support classroom teachers in all 15 elementary schools. At the same time, the materials center was established, and professional development institutes were designed and put in place during the first year of funding. This strategy of "starting large" was not restricted to small districts, such as Benton. The second phase in Bolton, with 60 elementary schools, took a similar approach with its LSC funds and provided training to 1,080 elementary teachers each year for four years. Though this occurred many years after Bolton's initial program establishment, it represented, in some sense, a re-birth or revitalization of the program and a re-visiting of the establishment phase.

As mentioned earlier, the lines of demarcation between phases are difficult to draw with precision. The end of the establishment phase is hard to pinpoint because much of the work is never "done," and the components of a hands-on program develop and solidify at different paces. For example, professional development strategies in both Lakeville and Benton evolved quite differently from the development of their materials centers. Professional development approaches evolved and shifted as the needs of teachers changed, as teacher leaders emerged, and as turnover created new demands. The continuing development of the materials centers, on the other hand, moved at a more consistent pace.

The next developmental phase of a program is *maturation*. Here, the focus is 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. On the whole, the program has been established and accepted; the materials system functions fairly smoothly; teachers are familiar with the kits; and professional development for teachers is available (most often on a voluntary basis), at least for teachers new to the district. Program leaders must work to continually assess the nature of the program's imperfections and endeavor to address them. This means that they must have the capacity to assess the program's status and make the needed improvements. Moreover, they 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.

Sycamore provides a good example of the nature of the *maturation* phase. Four years after the kit-based program had formally begun, it was established in all grades in all elementary schools. A centralized materials center was in place, and kit training was being provided. At that time, leaders initiated two important professional development strategies. The first involved the development of peer coaches. These were classroom teachers who had received about 100 hours of training at intensive NSF-funded summer institutes and were then available to provide their colleagues with coaching and support for their use of science kits. In contrast to the kit training that introduced teachers to the new curriculum, the addition of the peer coach addressed the program's longer-term needs of building capacity for professional development and leadership. Peer coaches were one way to help maintain an understanding of the status of the program and to ensure that the program was taught across classrooms by providing teachers with help and guidance. The second set of strategies involved the leaders' participation in a series of conferences and national initiatives that centered on supporting hands-on science programs. These activities connected Sycamore's leaders to others across the country who were engaged in the same enterprise, thus providing them with an expanded perspective on inquiry science programs and valuable support for advancing their vision.

Both of these professional development activities, although aimed at different audiences, were similar in that they reached beyond the immediate need of circulating kits. Rather, they targeted longer-term issues of building the capacity to continually assess and support the program across the district. These are the kinds of activities one finds in maturation.

The third phase of development is *evolution*. The hallmarks of the evolution phase are growth and improvement. These programs are still not immune to threats, nor are they perfect in structure or design, so leaders in this phase work to keep these issues in check while, at the same time, constructing for themselves a more expanded and somewhat 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 knowledge about 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.

Looking once again to Bolton, one can see that the science program has expanded and contracted over 30 years. If one considers its long history, it is possible to see periods when the strength of its basic components, such as professional development or the materials center, have waxed and waned. However, it is also possible to see clearly the steady maturation of the very ideas about hands-on science that have guided the program over time. One of the hallmarks of the evolution phase—interest in developing a richer understanding of science content and pedagogy—is quite evident in several ways.

First, prior to receiving their NSF Local Systemic Change (LSC) grant, program leaders undertook a large and sophisticated restructuring of the science curriculum. Throughout this process, the core team relied on the national science standards and their ability to articulate their learning goals to guide the district's work. This was a far cry from the early days, when activities and available materials were enthusiastically assembled in tubs with little thought for how they would work together within and across grades. What had changed was the development over time of a more cohesive understanding of science teaching and learning, as well as a richer understanding of what inquiry science really means. Bolton leaders have pushed themselves to understand inquiry more deeply and take their own professional development in this regard very seriously. As their own understanding grows, they look for ways to pass it on to classroom teachers. As they design professional development programs, they continue to emphasize and increase teachers' exposure to real inquiry and advance teachers' implementation of kits beyond "mechanical" use.

As the descriptions of the stages of development suggest, programs never shed entirely the threats and challenges of earlier phases. Rather, they continue to address ongoing issues as they take on a new set of goals associated with their continuing development. Moreover, with each additional set of goals, there are important implications at all levels of a school system: at the classroom level, the school level, and the district level. For example, leaders' desire to develop a leadership team with an expanded understanding of inquiry teaching may require actions at the school and district level, while their desire to foster the integration of science with reading might require actions at the classroom, school, and district level. To be sustained, program goals must be realized at different levels, which require 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 the factors identified in this study (described in Part III) at any of these different levels of the system. Together, the phase of development and the program leaders' level of orientation determine the factors' importance and priority.

PART III: FINDINGS

INTRODUCTION

The stories of the elementary science programs in this study are complex. Many factors have contributed to and inhibited their sustainability over time. Factors identified across the multiple sites in the study fall into three general categories:

1) Factors that pertain to the surrounding conditions, which describe the influences of the context in which the program operates
2) Factors that pertain to the science program components, which describe the role that concrete elements of the science programs (e.g., curriculum, professional development, leadership) have in contributing to or inhibiting sustainability
3) Factors that pertain to the whole science program, which describe overarching contributors to and inhibitors of sustainability that affect the program in less tangible but powerful ways

These factors do not operate in isolation. They interact with each other and shift in importance and influence over time. And, as with all complex stories, the factors are often intertwined and difficult to distinguish one from the other. To examine each one, it is necessary to draw what sometimes may be arbitrary distinctions between them, but the web-like relationships should not be masked. Rather, their complexity is a finding in and of itself. It explicates the range of pressures that come to bear on the sustainability of a program and the difficulty program leaders face in anticipating or controlling for them. In this vein, we have made these relationships explicit by pointing the reader to sections where associated ideas are discussed.

An in-depth discussion of each of 15 key factors follows. For detailed descriptions of how these factors emerged in the science programs, see the individual site reports.

SECTION 1: FACTORS THAT PERTAIN TO SURROUNDING CONDITIONS

SCHOOL SYSTEM CULTURE

- *A shared culture of collaboration and respect can support the establishment, growth, and evolution of sustained programs, while a competitive culture that illuminates rivalries can inhibit them.*
- *Even when there is individual will and interest, a district culture that lacks established communication avenues can stand in the way of taking actions to support a sustained program.*

- *Tensions between centralized services and a decentralized district culture can negatively affect sustained programs.*
- *A district culture that promotes learning and outreach can benefit sustained programs.*

Culture can have many meanings. In this project, it refers to the nature of the human, structural, and systemic environment in which the science programs functioned. 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 expected practices (e.g., volunteerism, support for professional growth, and extent of support for innovation).

Though project leaders might define culture in a far less formal way—perhaps by simply saying, "It's the way we do things around here"—they do not underestimate its power. They work within its confines and recognize that it affects the ways their goals, strategies, and communications are interpreted. As an influential condition, culture sets a foundation for the ways and extent to which the other factors described below contribute to and inhibit the sustained program. Thus, efforts to bring a science program to fruition must be compatible with the culture or, even though well intentioned, they are likely to fail.

The most obvious attribute of culture in the districts included in this study was the extent to which there were avenues for regular, productive communication. In Sycamore, for example, the long lasting stability in the central office and school staff has created an exceptional environment that has been fertile for the growth of a program. Sycamore teachers and administrators understand one another's styles and know how to communicate and collaborate with forthrightness and patience. As a result, they were far less concerned about building and maintaining personal empires within the district's hierarchy and much more interested in focusing on program goals and developing sound strategies for reaching them. However, more recently, with a large contingent of new teachers, the high level of collaboration and communication has not been as obvious.

Hudson is another district where individuals knew one another well; they have created an environment of trust, collaboration, and respect. Like Sycamore, this was, in part, an outgrowth of the fact that the district staff has been relatively stable over the years and, as a result, they have enjoyed the continuity and familiarity of personnel. In both cases, program leaders, Allison Stowe in Sycamore and Linda Lawson in Hudson, had been able to act with a degree of autonomy, understanding that their district colleagues trusted them to make appropriate decisions and would support them along the way.

Other district leaders, however, face struggles as a result of the district culture. Garden City, for example, has a well-developed bureaucracy that controls the way business is done. Communication is guided by an awareness of turf, and there is a heightened awareness of the need to protect one's own and respect others' areas of oversight. Thus, implementing a reform effort, such as a hands-on science program, can and does exacerbate the difficulties in communication that already exist. In this environment, science coordinator Fran Reece came to rely on the unique, strong personal relationships she had developed with individual district leaders, principals, and teachers who could support her in the face of a challenge. In this way, she was able to circumvent some, but not all, of the pitfalls that the district's culture presented.

Glenwood, on the other hand, offers an interesting example of how the culture of communication and collaboration varies within a single program. Generally speaking, Glenwood is an entrepreneurial place where district leaders nurture and even encourage competitiveness through their focus on obtaining external funding. Such efforts can and do illuminate rivalries between departments and between the school district and the external entities involved in particular initiatives. This is easily seen in the difficult start to the collaboration between Glenwood, the local science museum, and the local state university, which eventually became the foundation for the Glenwood program (see the discussion on partnerships). In contrast, within the science program itself, program leaders Sondra Calder and Judy Larson are deeply committed to a collaborative approach to project leadership. They are inclusive in their planning and decision making and act in ways that recognize and address the importance of broad-based participation.

The Lakeville program also experienced variation in communication and collaboration, dependent in part on where it was in its development and changing circumstances. For example, the collaboration between the district and the university struggled at first due to differences in culture and communication styles of the leaders. Eventually, however, leaders at each institution found ways to work with one another and successfully collaborated even under the most trying conditions. The relationship between the science program and the larger district, however, did not have as favorable of an outcome. During establishment and into part of the maturation phases, communication (and resultant district support) was solid. However, this was primarily due to the commitment of the assistant superintendent and his commitment to the program. Eventually, the assistant superintendent departed, leaving the relationships between the district and the program vulnerable. Despite the early support, the district did not have a larger culture of collaboration and communication, and Program Leader Paige Wolters and her university partners were left with the frustrating work of advocating for the program in the face of skepticism and increased accountability, sometimes resulting in tension and resentment.

In addition to the general pervasive culture of communication and interaction, the programs in some communities have been affected by a tension between the centralization of services as evidenced in the science program, and a culture of decentralization as evidenced in dispersion of management, decision making, and finances. In these places, even when program leaders conceive of appropriate staff development and instructional materials, they are beholden to the interests of the school principals and staffs. This is discussed more in depth in the section on decision making (p. 28), but is highlighted here because, in some cases, districts with a culture committed to the independence and autonomy of decentralization pose some particular challenges to the sustained programs.

For example, Garden City is a very large district, and its history as a consolidation of 82 once separate school districts reinforces the tendency of people to think and behave independently. Though more than 50 years have passed since the district formed, there remains a sense of competition and a lack of trust that is exacerbated by the fact that, in this district of nearly 800 square miles, schools are physically quite distant from one another. Thus, establishing and growing a districtwide program that, by definition, requires some degree of conformity is even more difficult than it would be in a place with more centralized operations.

The challenge of decentralization is not, however, limited to large districts. Benton, in contrast to Garden City, is very small (only 15 square miles), but it, too, has a culture of school autonomy and independent decision making. In part, this cultural characteristic is a reflection of a community that is highly political and, in the words of the superintendent, "very process-driven." Parents and community members watch the workings of the school system, which are tied to city politics, closely. As a result, science coordinator Constance Conner has needed to implement the science program with careful navigation of administrator and city stakeholders' personal and political agendas.

In addition to larger issues of needing to develop compelling reasons for why schools should participate in the program, decentralized environments also pose practical challenges for the science program. The independent nature of the individual schools in Benton, for example, has an impact on professional development activities of the district resource teachers. The resource teachers' work is shaped by the interests of each particular school and, at times, has been dominated by responding to individual teacher requests. This stands in contrast to a place such as Bolton, where resource teachers' work was more coherent and uniform across the district. In fact, at one point, a Bolton resource teacher was so proactive, her actions were perceived as evaluative, causing the union to step in and take action.

Another aspect of district culture—the extent to which a community typically reaches out to external sources of knowledge and support—also influenced the sustained programs. Generally speaking, outreach to individual colleagues in other districts and national organizations benefited the programs enormously. Sycamore, for example, benefited greatly from the input of several outside sources. Science Coordinator Allison Stowe collaborated with the local university, locally and nationally recognized science museums, an NSF-supported curriculum dissemination center, and an NSF- supported network of school districts engaged in elementary science reform. She commented on the benefits of these interactions and their influences on her decision making.

This kind of information seeking is supported by Sycamore and is an accepted part of daily operation. In contrast, the culture of Garden City places little value on going "outside" for professional development and consultation. While supporting the development of resources close to home, Garden City does not benefit from the knowledge and experiences of others. Science Coordinator Fran Reece uses her own vacation time and money to participate in experiences she believes will benefit the program through the gain of practical information and her own professional development. Thus, in this endeavor where learning curves are large, one can speculate that the insular culture has hindered capacity for growth and evolution of the program.

The examples offered above illustrate that culture is a backdrop that influences program leaders actions and the interactions of the factors that support and inhibit sustainability. Readers will be able to draw links between culture and some of those factors including decision making, leadership, implementation, money, and adaptation (to name a few) with ease. Simply put, culture is pervasive and, though at times difficult to accurately describe or interpret, a key influence on the district operations surrounding the establishment, growth, and evolution of the hands-on programs in this study.

DECISION MAKING AND POWER

- *Program leaders have little formal decision-making power or authority over the elements of their science programs.*
- *Decisions are made at many levels in a district by many different stakeholders. Any single decision can advance or inhibit the status of the science program.*
- *Leaders of sustained programs must find ways to navigate the decision-making structures in their districts and gain access to those who have the power to influence the status of their programs.*
- *The support of the central office is critical to the well-being of the science programs.*

Leaders have relatively little control over the many pressures and issues that can and do influence the growth and development of districtwide hands-on science programs. On the contrary, there is a wide range of decision makers in a district, each with his or her own allotment of formal and informal power, who can advance or inhibit a science program's growth and development at any point in time. Given this relative lack of control, leaders of sustained programs must understand their district's power structure and be adept at negotiating it in order to exercise what influence they can over the decisions that will affect their programs.

Each of the districts in this study has its own process for making policy and program decisions, some more explicit than others. Glenwood is characterized by its collaborations, formal alliances, and inevitable conflicts among partners with competing interests. Garden City is more hierarchical, with a formal organizational structure in place that is paralleled by an implicit set of rules and turf boundaries dictating who has access to those with the "real" power. Both Sycamore and Hudson, where superintendents, principals, teachers, and others have been in the district for decades, have stability that supports an environment where decisions are made based on the trust and understanding that has built over time between colleagues who have both work and community in common.

These different styles of making decisions and allocating power form the landscape within which science program leaders try to advance their programs. In doing so, they attempt to shape all of the elements of their program; however, the degree to which they can actually influence these elements varies widely. Some elements, such as the curriculum itself, professional development, and materials management are more easily contained within their purview. But others, such as budgets, state frameworks, and high stakes assessments, are critical to the program's well-being but far beyond the reach of program leaders. Also important in the program leader's landscape are the many different levels in a district at which decisions are made about whether and how the science program will be implemented. Figure 2, Hands-On Science: The Levels of Decision Making, illustrates these levels and the stakeholders present at each.

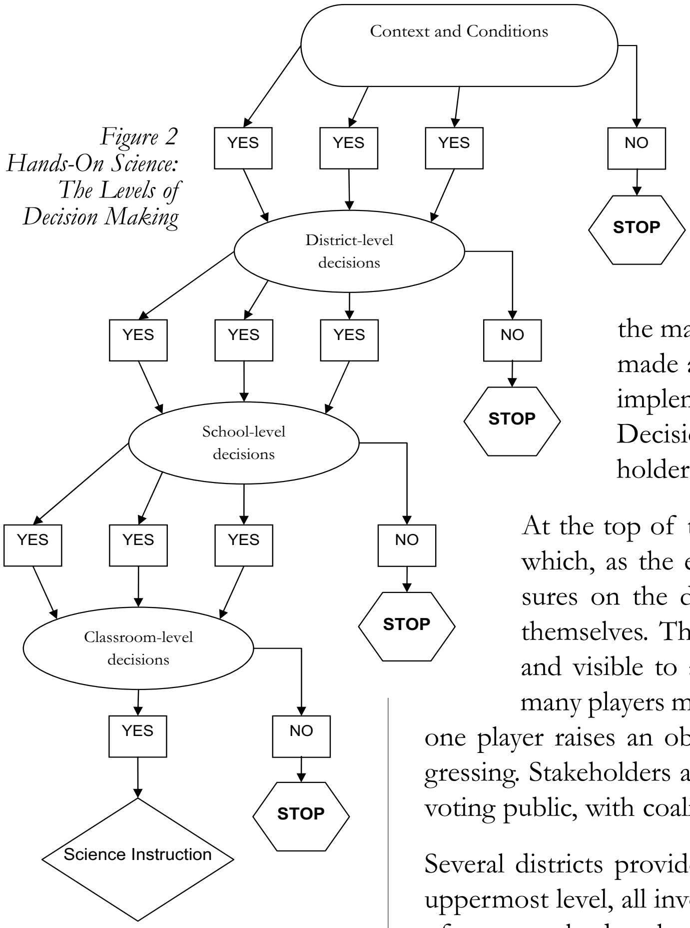


Figure 2
Hands-On Science:
The Levels of
Decision Making

At the top of the tree are the community contexts and conditions, which, as the earlier discussion pointed out, exert their own pressures on the district and, by extension, on the science programs themselves. These decisions are, almost by definition, quite public and visible to all. At that level, as at the district and school level, many players must concur for the program to move forward. If even one player raises an obstacle, that player can stop the program from progressing. Stakeholders at this level are generally the broader community and voting public, with coalitions, factions, and a range of perspectives at play.

Several districts provide striking examples of how complex events at this uppermost level, all involving many stakeholders and all beyond the control of program leaders, had a major impact on their programs. In Lakeville, for example, a conflict between the state and district science standards escalated into a heated debate among the school board and community members about the quality of the program. Opinions on the school board wavered in the face of public accusations of poor quality and conflicts of interest. The program would have been dissolved almost literally overnight had it not been for an impressive show of support from teachers, principals, community members, and parents. In Montview, on the other hand, the program was not so fortunate. The public's concern about the district's inefficient use of funds and its support for trimming back all centralized administrative functions resulted in the elimination of the centralized support structures that had enabled the science program to function successfully. Without centralized leadership, professional development, and materials support, the program appeared to wither on the vine.

A similar pattern holds true at the district level, where many constituents play a role in the ongoing status of a hands-on science program. Program leaders advocate for the needs of their programs by seeking the explicit or tacit approval from the superintendent and assistant superintendents, the school board, union leadership, and/or other curriculum leaders. Decisions at this level might be less conspicuous, but they are just as meaningful and equally as capable of stopping the program in its tracks.

Although programs do not require immediate access to the central office administration to survive, they do need to secure general district administrative support, or they will be in jeopardy. Superintendents' and others' decisions to attend to the needs of the science program, be they concerns about resources or other issues, have a cumulative effect over time. Continuing support sends a message to others in the district that science is important, just as a lack of interest affirms the opposite view. Merely looking at the most concrete example—science program finances—illustrates the core nature of this required support. Science program budgets undergo considerable scrutiny and depend on the support of the superintendent, the board, and central office administrators to maintain a reliable source of funds. Each year of regular, consistent financial support helps to embed the program more deeply as accepted practice. Clearly, gaining access to those with power and decision-making authority is key.

In Lakeville's earlier years, for example, the program had the close support of the assistant superintendent, which was critical for the science program's establishment. In him, Lakeville's program leaders had an eager and skillful advocate who was predisposed to ensuring that the program was as well-provided for as possible. Similarly, in Sycamore, the superintendent who arrived in the middle of the program's establishment strongly supported the science program. He had personally experienced the early roots of hands-on science with *SCIS* kits in the district years before and was philosophically committed to student-centered approaches due to his work with their Follow-Through program. In Bayview, the program's first leader, John Evers, paid vigilant attention to the program and district budget, and early on, cultivated relationships with his superintendents, carving out for himself a "seat at the decision-making table." When he was promoted to assistant superintendent, he was well-prepared to attend to the program's needs and understood them in the context of the larger educational program. And in Garden City, Fran Reece made strategic investments in access by inviting central office administrators to join her team when they visited "exemplary sites" and a science symposium. One of those administrators would eventually become the superintendent who, until the time he left, was an accessible support for the program—not a small feat in the bureaucratic Garden City environment.

As programs evolve, the pathways to access and power sometimes need to be reinvented. All of the programs in this study experienced changes in district leadership, and many suffered from a loss of equilibrium while program leaders worked to establish relationships with new central office administrators. Garden City and Sycamore, for example, most recently brought on new superintendents. This was a great loss in both communities since both Fran Reece in Garden City and Allison Stowe in Sycamore had long-standing collegial relationships with their superintendents. Additionally, their challenge was complicated by the fact that both superintendents came to their districts from somewhere else. After years of leadership from a long-standing member of the school district community, educating the new superintendent was, to some extent, like starting from scratch. Until new superintendents come to know, understand, and value the science program, its budget is in question and its future uncertain. For these reasons, districts where leadership has been stable, such as Hudson, have had a distinct advantage over more volatile places where superintendents have come and gone, such as Lakeville and Glenwood. In these places, leaders must continually re-establish the value of their program, taking valuable time and energy away from the many other tasks they must perform.

While Bayview and Lakeville were fortunate to have had a close "friend at court," this was not the case with most leaders, and, by necessity, they found other ways to promote their program. In Benton, for example, where program leader Constance Connor did not have such a strong advocate in the central administration, she worked to strengthen and elevate the role of science in the overall educational program by bolstering the relationship between science and literacy. She led an extensive effort to include the use of science notebooks in the curriculum and accompanied it with a teacher's guide and professional development. Another approach can be found in the early years of Bolton's and Hudson's programs, where the two leaders, Pearl North and Linda Lawson, assumed a considerable amount of informal power and influence by virtue of their charismatic personalities, dedication, and strength of vision. As their reputations grew beyond the borders of their districts, so did their influence at home. They were well-equipped to gain access to their superintendents and school boards and build support for their programs through these relationships.

North and Lawson amassed a significant amount of influence and, with it, were able to secure the resources they needed for their programs from within their districts. Other leaders, recognizing the volatility of shifting resources, changing priorities, and administrative leadership, sought a larger measure of control of their programs through access to external funds. Glenwood, for example, has consistently been successful at securing large grants: over $18 million in the past 12 years. These resources have played an important role in the dynamics of the district, providing a degree of independence for the program while, at the same time, creating opportunities for both collaboration as well as competition. Lakeville, Benton, Bolton, and Hudson also took advantage of the opportunity to secure large NSF grants. Although temporary, the money provided all of the leaders with the resources necessary to hasten their program's advances in professional development, curriculum, and materials management and with the independence to implement those advances as they saw fit.

Even when program leaders have access to key decision makers at the community and district levels, the ultimate decision about delivery of the program rests in the schools. Program leaders have no formal authority to influence the actions of either principals or teachers, particularly in districts that are formally decentralized. Rather, mandates from the district office and the influence of parents hold sway. And in the cases where principals have formal authority, program leaders need to be able to generate compelling arguments for why schools should invest their dollars in the science program. Ultimately, principals' attention to science informs the school staff that science should be taught, while its absence from the agenda informs teachers that it may be avoided. It is noteworthy that a great majority of principals who responded to the RSR survey (72–89 percent) reported that they "strongly agree" that they actively support teachers' efforts to teach science. Responses on a corresponding question on the teachers' survey, however, demonstrated that far fewer teachers (23–49 percent) strongly agreed their school administrations actively supported science. (An exception to these figures was in one site where only 50 percent of the principals reported strongly supported science with a corresponding 14 percent of their teachers saying the same). These figures suggest that even when principals believe they are supporting the science program, their actions may still inadequately convey those views to their teachers and the principals, ultimately, fail to adequately support program implementation.

Given the authority that principals have, program leaders of sustained programs must rely on their abilities to form relationships with them and other school-based personnel. Most of the program leaders have been able to build just such a rapport, and they are often viewed as extremely accessible and credible. For leaders who came from the "ranks" of classroom teachers, such as Lawson from Hudson, Wolters from Lakeville, Harder and Clay from Bolton, and Evers and Cooper from Bayview, their credibility was built in. Others, such as Fran Reece in Garden City and Constance Connor in Benton, had to develop that trust and respect over time by paying steady attention to teacher and principal needs.

The ultimate decision makers regarding science instruction are the classroom teachers. Although they are influenced by all of the factors that have been discussed and many more that will follow in this report, they determine what takes place in the classroom. Particularly in view of the minimal accountability for teaching science in the districts in this study, as well as the increased attention to student achievement in reading and math, there is little to compel teachers to provide science to their students except their commitment to offering a well-rounded education. Here again, program leaders must rely on their abilities to interest and engage teachers in the science program. In the absence of formal power and decision-making authority, they have few other tools with which to accomplish their goals.

SCIENCE FOR ALL

- *A centralized (or districtwide) program is considered an equalizer for schools and students, who may otherwise experience inequitable distributions of resources and variable classroom experiences.*
- *In the absence of accountability, equity suffers.*
- *Given the equalizing nature of a districtwide science program, when equity is expressed as a goal and value of the district, that goal isn't necessarily translated to support for the science program.*

"Science for All" 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"—the issue of equity—emerged as a factor in the sustainability of the science programs in three main areas. The first area focused on whether all students had access to science instruction, the second area on whether the science program was delivered equitably, and the third area on the value of the science program for specific populations.

Nearly all of the districts in the study had schools with widely varied access to resources. This was primarily due to variation in the economic support of the schools' surrounding neighborhoods, with those in more affluent areas benefiting from increased parent involvement and support in the form of volunteers, partnerships, and fundraising. For example, teachers in Glenwood felt that schools in low socio-economic status areas were not able to get the same resources that schools in areas with more active parental involvement could. The situation was similar in Garden City, where one school's neighborhood PTA raised enough money to fully stock the science laboratory while other schools had only minimal parental involvement. Even in Benton, where the various neighborhood populations are somewhat more consistent, inequities emerge in spite of the fact that Benton has a school choice plan intended to ensure that the school populations reflect the city of Benton as a whole. As one Benton principal commented, the plan "allows segregation of a certain extent to occur."

In this environment, teachers and administrators often recognized the science program as an equalizer, particularly in places that generally operated in a decentralized manner; such as Benton, Montview, and Garden City. In Benton, for example, where some described the school choice plan as fueling competition between schools, they believed that the science program helped to level some of the differences because all schools had access to the same program with the same materials. Similarly, administrators in Garden City felt that the increased movement the district was making toward centralization, as evidenced in the science program, was an effort to address the very visible differences in resources between the schools.

The appreciation for the science program as an "equalizer" also emerged in those sites that suffered less from differences in resources and more from inconsistencies in curriculum. When the Bolton program began, for example, Pearl North found that elementary science in that community was a "hodge podge" of science teaching, with some students receiving no science at all. Similarly, Thomas Donahue's work in Montview grew out of a desire on the part of district leadership to "unify the various community and school differences." In these cases, the science program was valued as a base of consistency.

Several districts' documents make explicit statements that express their interests in providing a strong science program for all students. In Glenwood, a science program description remarks: "All elementary students…will receive a quality science education program…." In Bayview, their "vision for science education" document states that they offered a "world-class, student-focused, K–12 science program that provides all students with the wide-ranging learning experiences…to solve practical problems, to inform decisions, and to learn more about taking responsible actions in their lives." Other districts that don't have statements explicitly targeting science still articulate more general equity-related statements in their district documents. Garden City's district plan for success, for example, articulates an "all students can learn" philosophy as does Sycamore's "Blueprint for District Improvement." The sentiment is particularly strong in Sycamore, where several administrators referred to a shift in community sentiment away from "good enough" toward seeking the best for all students.

Generally speaking, all of the districts in the study have systems that ensure all schools have access to the science program; in fact, this is one of the features that defines the programs as districtwide. Further, all of the program leaders have worked to implement strategies to ensure that all teachers and principals receive sufficient training and support to use the program. Professional development plans offer training and in-depth professional development opportunities to all teachers and, in several cases, they are required. Districts also train and support individuals to provide customized support to teachers, whether in the form of peer coaches in Sycamore, science staff development teachers in Benton, or resource teachers in Lakeville. Some programs, such as Bolton's involvement with the *Kagan Cooperative Learning Project* and Glenwood's case study teams, even focus specifically on approaches that will broaden student participation. And yet, despite the program leaders' best efforts to ensure equitable program delivery, the data shows that program use within each district has been highly variable.

The study found no evidence in any site of a system for ensuring that the science program will, in fact, be taught. Some small mechanisms, such as monitoring the materials used in kits or including science instruction in teacher evaluations, have been implemented by individual principals committed to the science program and to ensuring that their students have equal opportunities to learn science. But these efforts are idiosyncratic. There are no districtwide systems in place to assess whether or not teachers are actually teaching science, and there are no districtwide consequences for teachers who fail to do so. The lack of accountability mechanisms is exacerbated by the increased accountability for literacy and mathematics. Not only is there little extrinsic incentive to teach science, but, in fact, teachers in nearly all districts face the disincentive of needing to give increased attention to reading and mathematics instruction. The end result, then, is that instruction is left to the discretion of the teacher, resulting in inconsistent and, by definition, inequitable instruction.

Relatively few (13–24 percent) of the teachers responding to the RSR survey reported that they teach the district-provided kits from start to finish as district leaders hoped and expected. Some of the responding teachers (7–38 percent) reported that they teach the kits from the beginning until they run out of time. But the majority of teachers (48–73 percent) reported that when they use science kits, they "pick and choose parts to teach." This reveals not only the variability of instruction in a district but also the extent to which that variability is entirely at the teachers' discretion.

The exceptions, of course, are those districts facing the arrival of high stakes tests in science. To some extent, the presence of a state test has increased the visibility of science and, subsequently, supported program delivery. Teachers in all districts that have a current or upcoming state test (including Bayview, Benton, Sycamore, and Garden City) reported being apprehensive about the arrival of their states' science tests, but only some felt that the test could have a detrimental effect on the science program. In fact, some felt that the impending arrival of a science test had bolstered the program and ensured that more teachers are using the curriculum. As one Benton teacher stated, teachers are "teaching to the test, but not necessarily in a bad way." Questions of appropriate alignment between the program, the standards, and the test are discussed more in-depth in the discussion of accountability that follows. Similarly, an in-depth discussion of the variability of the curriculum can be found in the discussion of quality.

By and large, equity did not readily enter into most of the conversations researchers had with teachers and administrators. The exception was in references that teachers and administrators in several districts made regarding their beliefs that the program has benefited students of particular populations. In Sycamore, for example, the director of special education praised the program for engaging students who otherwise might have been left behind in science instruction. In Lakeville, the science program is highly valued by many teachers, but particularly by those working with students who are English Language Learners. Similar sentiments were expressed in Glenwood and Bolton, underscoring the potential of the hands-on science programs to more effectively meet the learning needs of all students.

Generally speaking, issues related to equity are elusive, manifesting themselves in restrained, subtle, and, sometimes, unspoken ways. One teacher, for example, expressed her resentment that resources were being directed toward establishing a special program in a single school, implying that the resources should be directed to more centralized needs, such as the science program. A principal in a different district articulated her discomfort with the fact that she has no extra resources to devote to science while acknowledging that other schools do, in fact, have access to such resources via their local neighborhoods and connections to community partners. The comments weren't limited to the school level; one of the science coordinators described the attention given to a new initiative in another subject area, suggesting that resources that could have been devoted to science were being redirected. Indeed, though many recognized the districtwide science program's potential, not only to provide science instruction to all students but also to contribute to making progress toward improved equity across the district, the interest in supporting this potential was never clearly articulated either verbally or in writing in any of the data collected.

SECTION 2: FACTORS THAT PERTAIN TO SCIENCE PROGRAM COMPONENTS

ACCOUNTABILITY

- *There is limited accountability for student learning or for the delivery of the program. This can either contribute to or inhibit the sustainability of the science program depending on the district context.*
- *In the presence of high visibility and high stakes tests, science is often overshadowed and, therefore, time and resources devoted to its accountability are diminished.*
- *When an accountability strategy for student learning or program delivery does exist, resulting data are of little use to program leaders if they have no power or authority to make and follow through on decisions based on that data.*

Two types of accountability have played a role in the sustainability of the hands-on science programs in this study: accountability for student learning and accountability for principals' and teachers' program delivery. Accountability measures for student learning include student written and performance tests, student work, and writing in student science notebooks. Accountability measures for program delivery, on the other hand, include requirements for principal observations of science instruction, tracking of kit usage, and analysis of school improvement plans. Generally speaking, some districts have district or state tests in place that provide the only mechanism for accountability for student learning. Mechanisms supporting program delivery, however, are universally weak. The presence and absence of these mechanisms, depending on the site and its context, sometimes support the sustained program and sometimes hinder it, but always cause high levels of concern and anxiety. The timing of this study, falling during an increased national emphasis on accountability, particularly for reading and mathematics, highlights the potentially enormous influence of a single factor and the fragility of even the most enduring sustained programs.

Any discussion of accountability must necessarily be contextualized by answering the question: Accountability for what? In the sites participating in the study, the answer to that question always was "standards." In most cases, the state standards have had the greatest influence, though there were some exceptions, such as Lakeville, Montview, and Glenwood, where district standards hold more, or at least equivalent, sway. Until the relatively recent arrival of the national emphasis on standards and their aligned tests, the attention to standards had been one aspect of local style and culture—some places (such as Benton or Bayview) operated with little attention to guidelines from outside their own communities while other places paid close attention. Garden City, on the other hand, had emphasized the role of state standards for some time. There, every person interviewed, from central office administrators to novice teachers, concurred that the state standards drove the district's curriculum in every subject area. This is not surprising in today's national climate when attention to and concern for standards and their tests is pervasive.

With this said, it is noteworthy for this study that differences between state standards and the goals of the sustained programs have established pressures on the science programs. Some sites face an even more complicated issue in that they have local district standards that illuminate goals for student learning different from those articulated by the state. In Glenwood, for example, the district has both content and performance standards by grade-level ranges, whereas the state science standards are essentially content standards. Similarly, Montview has developed their own set of performance expectations that move far beyond the guidelines outlined by the state. Lakeville, on the other hand, has local standards that stand in direct conflict to the state. This example is described more below.

Accountability for Student Learning

Four of the nine districts in this study—Benton, Sycamore, Bolton, and Glenwood—have a formal student test in place to measure student achievement in science, while two districts—Garden City and Bayview—are looking at tests on the near horizon. In Benton and Sycamore, science was included in the state's assessment program. Bolton has included the science component of the California Achievement Test in its district's assessment schedule, and in Glenwood, the test is a district-designed assessment.

In those places that currently have science included in the state test, or are anticipating the addition of science in the future, anxiety about the results runs high. In most cases, the test results are highly public and the focus of much local discussion and concern. At the same time, without exception, no high stakes are attached to poor performance on the science portions of these tests. As a result, the impact of the presence of these tests on the sustained programs has been mixed. On the one hand, the presence of the test sends a strong message that science has an important place among the core subjects. And yet, the absence of consequences for poor student outcomes will dampen this effect significantly.

Indeed, in the four places where a test has been regularly administered, student science scores appear to be of little concern to those outside the science program leadership. One explanation for the lack of attention is the fact that, in nearly every site, tests that focused on literacy and mathematics are high stakes and have potentially serious consequences. Thus, even in those places with science tests, the importance of elementary science has been greatly diminished and concerns about accountability for it have been placed on the back burner. In those places with no science tests at all, accountability for student learning of science is not even on the proverbial radar screen.

In Glenwood, for example, district leaders have invested a great deal of resources into the development of a districtwide test. The test has three types of items: multiple choice, open-ended response, and performance tasks. The development of the test has been meticulous and ongoing, seeking to draw close connections between student growth and each of the district's content standards. Though fairly expensive, program leaders see a good deal of value in the test beyond the obvious source of student outcome information. They believe that use of performance tasks help to reinforce the district message about the importance of students actually doing science, not just reading about it. But still, even in light of the fact that this test can be considered a formal accountability system for science, the consequences for the test are relatively low stakes. In fact, some of the principals and teachers complained that the test is poorly timed and that the results are delivered too late.

In Benton, low scores on the state standardized test, though they have no formal consequences, are an embarrassment to the program. The district was hoping to develop new, authentic assessments (perhaps not very different from those in Glenwood) to offset the disappointing scores, recognizing that the absence of such tests make the program vulnerable (as illustrated by the Lakeville example below). Still, the fact that science was tested gives it much needed stature, especially in light of the priority on literacy. Though some felt it is a distraction, others felt that the test helps to raise the visibility and credibility of science. Some of the questions are, in fact, aligned with the goals of the science program, and one resource teacher commented that "some teachers are now teaching to the test…but not necessarily in a bad way."

Because student achievement in science has received little public or administrative attention, one might conclude that the presence or absence of student achievement data in science has been irrelevant to sustainability. This appeared to be the case in the study sites with one recent and important exception. 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 fully developed. This gap leaves the program in a vulnerable position. In 2000, some 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 to combat the challenge and, thus, preserved the program. But, the vulnerability of the program at that moment should not be underestimated.

Hudson illustrates the importance of context in determining the impact of accountability measures on the sustainability of a science program. After Hudson's program had been in place for well over 10 years and had received strong local support and national recognition, the district administered a science test to sixth graders to assess the impact of the program. Students' scores were not very high, but rather than question the strength of the program, the merits of the test were questioned. One administrator recalled, "We knew there was a mismatch there… The kids were learning more science this way than they were the other way. Sometimes you just have those gut- level feelings."

Whether or not students were learning more science with Hudson's program is less important here than the fact that it was possible at that time to disregard test scores and, with no other data at hand, rely on "gut feelings" to justify the quality of a program. In the current climate, which one teacher characterized as a "testing sickness of politicians," empirical data is an important tool to have at hand when the value of a science program is questioned. In the absence of these data, programs are fragile.

State- or districtwide tests notwithstanding, almost all of the programs have attended to developing assessments for teachers to use in their classrooms with each unit. Bolton has spent considerable time on this, developing assessments as well as professional development in their use for teachers. The same is true for Bayview, and Benton has developed the use of science journals to both assess student learning and tie science more closely to literacy in the district. Teachers' use of these assessments varies widely, however, and there is no district in which they are used systematically and where student outcomes are used to inform professional development or curriculum development.

Accountability for Program Delivery

In the nine districts involved in this study, there was no evidence of any formal mechanisms explicitly devoted to ensuring program delivery. There is no requirement, for example, that principals observe science instruction (such as commonly exists for reading and math) and, in fact, the program leaders themselves find it difficult to stay informed about the status of the program by means of direct observation. Not only are there simply too many classrooms and too few staff, but the cultures of teacher evaluation and accountability also are a hindrance. In Bolton, for example, program leaders' ability to observe instruction is hampered by the union concerns that they are exercising oversight authority they don't possess. Although most districts have requirements for the amount of time that should be devoted per week to each subject, including science, there are no mechanisms to ensure that this is being fulfilled and, as one principal in Garden City commented, "The only way to know is to be in that classroom."

The only exception to the widespread lack of information about program delivery is the fact that some programs gather information about kit use based on the status of the kits and teacher evaluation forms returned to the materials centers. In Hudson, for example, clerks generate annual reports that reflect the extent of kit use by each teacher, and distribute them to principals for their review and follow-up. However, the extent to which principals make use of these reports varies widely and is limited at best. Similarly, Glenwood uses teachers' requests for consumables as an indicator of implementation and provides this data to principals, but no principals reported that they use this information to guide planning for science or as part of teacher evaluation. In fact, no Glenwood principals reported ever having observed a science lesson to evaluate a teacher. The RSR teacher survey data corroborates this, demonstrating that a vast majority (76–92 percent, depending on the district) of teachers have never had their principals observe their science teaching.

In two cases, Glenwood and Benton, a potential source of data on program delivery exists, but it hasn't been incorporated into routine practice.

Specifically in Glenwood, since 2000, as a result of their NSF-funded USP[5], there have been specific goals for improvement in science, and the school site plans have had to include specific language reflecting those goals. Still, even though science has taken a step forward in that it was explicitly included as part of this school site plan, there are no specific sanctions for not meeting the science goals. In Benton, resource teachers provide principals with a form designed to facilitate their observations in science. The form lists the important concepts in each module and clearly describes evidence of student learning. However, only one principal interviewed reported using this form.

Although science tests are not included in most states' current high stakes accountability systems, even the anticipation of a test and anxieties about its potential impact has affected the sustained programs. In Bayview, for example, leaders are examining their curricula to make sure it is in alignment with the standards and, presumably, the coming test. As a result, not only have the units been shifted by grade level or replaced, but the program has moved from one essentially driven by teacher choice to one that prescribes which units will be taught and when. Many teachers have found this to be a difficult adjustment, but understand that these steps had to be taken to best ensure that their students will perform well on the test when it arrives.

Benton also has had concerns about alignment with the state standards and test. With tests in Benton highly publicized and the superintendent stressing the importance of good performance (students must pass the reading and mathematics portions of the grade 10 version of the test to graduate), leaders of the science program are actively working to adjust the curriculum. One resource teacher made an item-by-item analysis of the science test and the Benton curriculum, saying "it brings to the fore whether the science curriculum is being delivered or not."

In this high pressure environment, the greater the lack of alignment between the test and the program, the more strongly program leaders have felt compelled to adapt the program. As one superintendent observed, "We can't afford to continue to develop a program we feel strongly about…if the state says, 'No, we are not going measure you that way.' If that is going to happen, then as much as we believe in this, we are going to have to change."

Another circumstance related to program delivery and science curriculum revision is the presence of accountability for literacy. In this national environment of increased accountability, no subject area is under more scrutiny than literacy. Not surprisingly, then, nearly all of the 194 teachers interviewed in this study suggested that they felt they had less time for classroom science instruction because of the increasing emphasis on reading. In

---

[5] The Urban Systemic Program (USP) in Science, Mathematics, and Technology Education is a program of the National Science Foundation. It is the key component of the division of Educational Systemic Reform's effort that supports urban school districts.

Hudson, for example, many teachers feel unable to include science regularly, if at all, because they feel the need to emphasize math and reading and, in some cases, teachers return kits to the materials center unused. In Garden City, where there has been a science test in place for many years, the recent absence of such a test during the re-vamping of the state testing system has reduced the importance of science. Most agree that science has taken a back seat to literacy and mathematics and that science is a "stepchild." A number of teachers went on to say that only standardized tests in science with high stakes attached will guarantee airtime in the classroom. In this environment, public reporting of school performance (as evidenced in all districts in the study), district budget increases tied to improved scores (as evidenced in Montview and Lakeville), and sanctions tied to poor performance (as evidenced in Glenwood and Bayview) together are contributing to a reduction in the quantity and the quality of science instruction.

In light of the strong emphasis on performance for literacy and mathematics, many districts have taken steps to incorporate elements of literacy into their programs. These strategies have a dual purpose. First, they provide teachers with ways to address the high priority of literacy, while at the same time giving more classroom time to science. Also, it is a way to "cover" more of the science content that, while included on the test, is simply too vast to address through the kits. In Bayview, for example, program leaders have made lists of literacy works that have science themes available. They are distributed with a caution that, though useful, teachers need to be careful not to make superficial linkages to science. Similarly, Sycamore has purchased readers to accompany its *FOSS*[7] kits and has worked to ensure that literature included in its "literacy initiative" is in support of the science program, meaning that the topics are aligned with those included in science instruction. About two-thirds of all teachers responding to the survey reported that they use science-related literature and nonfiction books either "often" or "very often."

Accountability is one of the clearest examples of how a single factor can either support or inhibit the sustainability of a program depending on the district's culture, context, and age. Accountability is a particularly prominent factor, perhaps because the study took place during the national advent of increased accountability and high stakes tests in all subjects. In Garden City, there is no doubt that a test will elevate science. One central office administrator remarked, "As long as there are scores to report, principals will be supportive of ways to support that content area." Program advocates look forward to the attention science will receive as a result of the test even as they acknowledge the skeptics who feel it will push the program in a different direction. One principal, for example, stated that the coming science test, "…is great…if we are going to value it…I can see it becoming a fourth 'R.'" Still another principal observed, "Teachers need to have faith that the kit is going to help with the CAT[6] in this age of accountability." However, the risk that it won't help is great, and the temptation for teachers to rely increasingly on textbooks in science is evident in Garden City and Lakeville, where recent adoption decisions have included textbook options.

In the face of demands for information on student learning, sustained programs can be vulnerable. Lakeville's experiences suggest that districts need to systematically collect, interpret, and communicate about concrete data. At the same time, however, most of the programs in the study thrived for many years with no such data. This suggests that, in the absence of specific accountability measures, program leaders and others make decisions based on limited and informal data sources combined with their own observations and perceptions (discussed more in Section 3 below) about the status of the program. Thus, a program can appear to be sustained—embedded in the system and accepted as standard practice—but not actually taught.

IMPLEMENTATION

- *Leaders of sustained programs have used a range of approaches to implementation with no single approach demonstrating more success than another.*
- *Central office support is a necessity for laying the groundwork and establishing the elements of a sustained program.*
- *Leaders of sustained programs choose implementation strategies that account for the culture of the district, district priorities, and the relative importance of the different elements of the program at a given time.*

Implementation refers to the strategies program leaders use to initiate hands-on science programs, and the methods they use to bring their science programs to be accepted as districtwide practice. Though all of the district leaders in the study have shared a similar challenge—establishing a program that includes resources, curriculum, professional development, and instructional materials—their overall approaches to implementing their programs have been highly variable. In Benton and Sycamore, for example, the leaders turn to outside consultants and resources for guidance on professional development design and how to best introduce and communicate about the program. Lakeville, on the other hand, collaborated with university partner scientists to devise a pilot school approach to implementation that would eventually scale up to work districtwide. Similarly, the Glenwood district leadership collaborates with two partners, each of which has a financial stake in and philosophical commitment to the hands-on program. And in Garden City, the leadership stands essentially on its own, with little external consultation or resources.

---

[6] Corona Achievement Test, Garden City's state standardized test.

It is worth noting that although each leader could have chosen to pursue any kind of science program, each chose to focus on hands-on science instruction. Whether their belief in the hands-on approach has come from exposure via a mentor or colleague, personal experience with hands-on instruction, or their own science background, all have been deeply committed to bringing the hands-on experience to their communities. The older programs also were influenced to some extent by the national political climate of the 1960s and 1970s that followed Sputnik and was concurrent with NSF's emphasis on developing science curricula and increasing the number of people pursuing careers as scientists. Two of the newer programs, Sycamore and Benton, also have roots in this era. During the 1970s, they both had small-scale programs that introduced teachers to hands-on materials using NSF-developed curricula. These programs were not districtwide, and there was a clear delineation between their existence and the establishment of the districtwide program, but they may have contributed to the fertile ground by setting a tone for acceptance of hands-on instruction.

With similar basic needs, the creative leaders of these programs have met them in a variety of ways that have been particularly appropriate to their situations. They each have shaped their own strategies for introducing and growing the reform in the district by taking into account many influences, including their own personalities and styles, resources available, restrictions attached to those resources, and district culture. Given the range of strategies that has worked for the districts in this study, one can conclude that no single approach to implementation necessarily leads to a sustained program.

Specifically, Benton used what one might refer to as a "snowplow" approach: With the support of an NSF Teacher Enhancement grant, the science coordinator set out to reach all teachers in all grades in a relatively short time by establishing summer institutes designed to provide large-scale training on the use of kits. Nearly one-third of all K–6 teachers attended the institutes in each of the four years they were offered. In contrast, the Hudson program has employed what one might refer to as a "wedge" approach: The effort to engage participants in the program was a "soft sell" and focused, at first, on engaging only those teachers who were interested in science, leaving those who were reluctant in the background until they could see the merits of the program as implemented by their more enthusiastic colleagues. Lakeville provides an example of a combination of these approaches. The district first established a single pilot school where it designed and refined a model that could be replicated in the other schools in the district. After three years of pilot-school work, the district supported expansion to five more schools and then, several years after the first pilot school was established, obtained funding from NSF and support from the Lakeville central office to expand the program to the entire district. Other strategies fall within the snowplow-wedge spectrum, such as those used by Glenwood, which began with a focus on developing the knowledge and expertise of groups of teachers, and by Garden City, where the program began by targeting a single grade level across the whole district.

Regardless of variations in strategy, all of the programs share a common characteristic in that, at the outset, they had the necessary strong support from the central office. Support could have been monetary, verbal, or both. In some but not all cases, the central office was the impetus for a program via a needs assessment or recognition that it was necessary to have a more coherent, consistent science program. In Benton, for example, the superintendent and school committee had committed to a five-year plan for improving the curriculum and implementation of every core subject, including science. This resulted in hiring Constance Connor as the science coordinator and supporting her efforts to establish the districtwide elementary science program. Similarly, in Hudson, the superintendent and director of mathematics and science called upon teacher Linda Lawson to take the reigns of science, usher it through an adoption, and bring it to the level of success of other recently revamped subject areas. In Montview, where the science program was initiated only a few years after the district itself was established, science coordinator Thomas Donahue was charged to develop a program that would "provide glue" to the various schools in the newly unified school district.

Later, as the programs moved from the establishment phase to maturation, implementation strategies and approaches shifted as well. Leaders' choices about implementation were not rigidly tied to the strategy that seemed most viable at the beginning. In fact, leaders demonstrated great flexibility as their strategic implementation choices reflected accommodations of their own and others' leadership styles, district and school cultures and changing district and community circumstances. Some of these choices focused on large-scale issues, such as how best to organize materials, how to maintain attention for science, and the best strategy for introducing a new component to the program. For example, program leaders in Bolton had to make a major shift in implementation strategy to accommodate unusually high teacher retirement rates in the district. Professional development training that had been conceptualized to focus on the challenging instructional issue of integration of science with other subjects had to be reconceptualized to accommodate the new-to-grade teachers.

Other implementation choices influenced by style, culture, and context came in the form of small-scale daily interactions between program leaders and individual teachers, colleagues, and administrators. While seemingly insignificant, these small actions were cumulatively quite influential on the sustainability of the program because they left lasting impressions that ultimately translated to positive program support. For example, in Garden City, the science coordinator operated within a cumbersome organizational structure that had few established channels for communication or collaboration. Still, she was generally able to reach the right people to help the program progress because she relied on her personal interactions with individuals in informal settings and capitalized on the trust and respect she had earned through her demonstrated commitment. Similarly, the science coordinator in Benton won respect and, perhaps even more important, loyalty to the program through subtle elements of the program implementation design. Her attention to personalized teacher experiences and in-depth learning in the professional development design did not go unnoticed and was much appreciated among teacher participants who were grateful to be "allowed to be learners."

In summary, an effort to capture a formula for implementation that would lead to a sustained program would be fruitless. Anyone seeking such a formula would be disappointed as these sustained programs all reflected highly variable and flexible strategies that shifted with district circumstances, funding, and leadership changes.

INSTRUCTIONAL MATERIALS

- *The curricula of sustained programs typically are composed of a combination of materials—ranging from homemade lessons to commercial units—and often have supplemental components which, in some cases, include textbooks.*
- *Instructional materials in sustained programs evolve and are adapted over time.*
- *A district materials management center provides symbolic and practical evidence that a hands-on science program has been sustained.*
- *Instructional materials for hands-on elementary science programs require processes and systems for development and selection; management, distribution, and storage; and acquisition and refurbishment that consume a great deal of human and financial resources.*

Instructional materials are an essential component of any science education program. All of the science programs in this study were primarily kit-based, meaning they were based on boxes that included a teacher's guide and the necessary manipulatives for teaching the lessons outlined in that guide. From the very start, program leaders in every site had a shared challenge—what materials to use; how to get those materials to the teachers; and subsequently, how to retrieve them and prepare them for the next teacher. While sharing similar concerns, they each devised a sensible, customized strategy given the financial resources, climates, and cultures of their districts.

All of the programs in the study fall somewhere on a spectrum ranging from developing materials locally to purchasing all materials from a commercial or outside source. Generally, districts with a long science program history, such as Bolton, Hudson, Montview, and Bayview, began with curricula that were either entirely locally developed or were combinations of locally developed materials and *ESS*[7] or *SCIS*[8] units. Other districts with

---

[7] *ESS, Elementary Science Study,* curriculum kits and materials were developed by Educational Services Incorporated (later to become Education Development Center (EDC) in Newton, MA. Development funded by NSF and begun in 1960.

[8] *SCIS, the Science Curriculum Improvement Study*, was founded at UC Berkeley in 1963 by Dr. Robert Karplus with funding from the National Science Foundation. The study developed science curriculum for levels K–8.

relatively young programs, such as Glenwood and Benton, established their programs with commercially available kits at the outset. And others, such as Sycamore and Lakeville, drew from noncommercial external sources, with Sycamore obtaining kits from the local science museum and Lakeville purchasing them from a model science program in Mesa, Arizona.

Garden City is an exception in that it has a program that was established more recently but used locally developed kits at the outset. After visiting two existing programs in the country, the science coordinator led an effort to create the district's own materials, using a nationally recognized program as a model (Bolton also used another district as a model for their locally developed curriculum). This was in 1990, when there were relatively few commercially available, comprehensive hands-on curricula available, outside of the materials resulting from the curriculum development efforts of the 1960s and 1970s. Garden City's program leader chose an approach that was compatible with the district's culture, which was somewhat insular and focused on internal strengths and capacity rather than seeking external input and expertise. Even though one teacher described this monumental task as a "haphazard" process that was "an ordeal," they made it through the writing process. After several trips to a discount store to buy materials, they had assembled their first kits together and were ready to distribute them to their first group of participating schools.

In contrast, the Benton program began in 1994, when the newly created NSF-supported materials were entering the market. This timing, combined with the fact that their initial establishment effort was funded by NSF, led science coordinator Constance Connor to assemble a curriculum that used only commercially available kits with a combination of *FOSS*, *Insights* and *STC*[9] units. This full-scale purchase and the organization of a materials center to handle storage and distribution complemented and supported Benton's large-scale implementation strategy.

Regardless of the ways in which materials were selected in the establishment phase, as the programs entered maturation, every district curriculum evolved and adapted over time. Looking back at Garden City, for example, as the program grew, Fran Reece began a process of upgrading and expanding the kits, replacing some of the teacher-developed materials with those that were commercially available. Her efforts were driven by her belief that the commercially available kits were of higher quality because they included more inquiry approaches and were more coherent. While one might think that teachers would object to the loss of the "home-grown" materials, they actually didn't distinguish a great deal between the two. They felt that hands-on science was their program, regardless of the origin of the materials they were using.

---

[9] *STC (Science and Technology for Children)*: Development by National Science Resources Center, published by Carolina Biological Supply Company.

Constance Connor in Benton also made adaptations to their materials. Having started with all commercially available units, their adaptations were driven by both external and internal forces. Responding to external pressures, she and teacher colleagues made adjustments so that the curriculum reflected changes in the state curriculum framework, the arrival of the statewide standardized test, and the requirements that specific topic areas be addressed in particular grades. Still, most of the curriculum adjustments resulted from internal concerns that the curriculum wasn't sufficiently coherent, and that the materials themselves either didn't work well or needed to be more customized to issues and approaches unique to Benton. Most recently, Benton had added a science notebook component to the curriculum, responding to both the external emphasis on literacy and the internally driven interest in having students exercise their observational, organizational, and descriptive skills.

Curriculum adaptation has been built into the Bayview program from the beginning. Long before most of the commercially published curricula were available, science coordinator John Evers led the local effort to write and gather materials for their units. His process entailed a constant system of writing, feedback, evaluation, and revision that continues even today. As the current science coordinator remarked, "The units of work are never finished." While this process has been a hallmark of the program for decades, like Garden City, Bayview also is introducing some commercially available kits in an effort to better align with state standards. Unlike Garden City, teachers have expressed criticism of those kits, feeling they are less substantial than and not as developmentally appropriate as those developed locally.

Most recently, adaptations to some of the curriculum programs have included the addition of texts. Garden City, Sycamore, and Lakeville all have incorporated texts into their curriculum primarily as a response to the increasingly prominent state requirements and accompanying high stakes tests. Given the fact that, generally speaking, textbooks are antithetical to the hands-on approach, these moves are troubling for those concerned about continuing the sustained programs. Until now, even when texts had remained in the classrooms, they were used only as references. The shift to include texts calls into question whether these programs will be able to maintain their core beliefs and values and continue to advance their programs.

Program leaders' decisions regarding what units to use have been shaped by their beliefs and values regarding hands-on science and quality instructional materials, and their understandings of the culture and accepted modes of operation for their districts. But, they have not had the luxury of focusing only on philosophies and strategic planning; simultaneously, they have had to give attention to the very practical issues relating to doing hands-on science. Each of them has had to plan for management and financial support for what is essentially a warehouse of teaching materials. Again, each has arrived at a customized approach, accounting for the resources and support available at the time. Ultimately, the materials management centers will provide critical practical and symbolic support for the sustained programs.

For example, the materials center and distribution system in Garden City was carefully planned and designed from early on. Reece took advantage of existing delivery systems and devoted time to educating herself about kit management through interaction with the Association for Science Materials Centers (ASMC). She was able to secure space from the district and support her very competent materials center staff through a range of funding sources, which were eventually absorbed into the centralized district budget. The system now functions so well, in fact, that Reece fears that some teachers are starting to take the science program for granted, forgetting that they have the unusual benefit of having materials for science delivered to them fully stocked and ready for instruction.

In contrast, some districts have established materials systems that are partially decentralized. Glenwood, for example, has so many teachers in a highly populated area that centralized storage and distribution of kits districtwide is impractical. Thus, their kits reside at the schools but are supported with a centralized system for refurbishing the consumable materials. This approach requires that someone at the school site assume responsibility for ensuring the kits are in the appropriate classrooms and have been properly restocked after each use. The Glenwood program leaders have tried several strategies for accomplishing this, including identifying teacher leaders at each school and hiring parents on a part-time basis. These strategies have seen variable success and, in general, each principal has needed to find a solution that will work in his or her school. This remains a vulnerable point for the program, particularly since the emphasis on literacy and mathematics has taken center stage and principals are placing their priorities in these areas.

Montview took a similar approach to Glenwood. In keeping with their decentralized culture, Montview leaders asked individual schools to design their own materials management systems that would ensure that teachers had the basic equipment and materials they needed. Schools seemed willing, perhaps, in part, because the science program leadership was highly respected and demonstrated an eagerness to help in any way they could. In some cases, science program leaders worked with the schools to flesh out the details of their materials management systems. Once the systems were in place, the district supplied centralized support for restoring and replenishing the consumables. They kept the supplies in a warehouse and provided teachers with a list of items they could order. In contrast to the practical issues shaping the management decision in Glenwood, Montview leaders consciously felt that a de-centralized approach would help protect the program from centralized cuts. They also felt that this approach, which required commitment and attention on the part of teachers and principals in the schools,would contribute to the sustainability of the program.

The eventual disbandment of the Montview central office shows that their anticipation of centralized cuts was justified. However, the resulting catastrophic effect on the science program calls their strategy into question. Where there have been centralized materials centers, even in the face of financial pressures and cuts in administration and professional development support, the materials centers have held steadfast.

Regardless of the design of the materials management system, district administrators frequently view them with pride. It is this component of the programs that is most often supported with a line item in district budgets, and the last component to suffer a threat of reduction or elimination. In many districts, this trend is sometimes a cause for friction between program leaders and central office administrators. In Lakeville, Bolton, and Benton, for example, when leaders advocating for increased funds to support professional development, improve the curriculum, or add staff, it is not uncommon for administrators to deny these additional requests and cite their continued funding for the materials center as evidence of their continuing support for science.

These findings suggest that, in fact, not only are materials centers necessary, practical supports for the science programs, but they also make an important symbolic contribution to the programs' sustainability. In some districts, the centers are viewed as a point of pride and perceived, to some extent, as evidence that the district is giving attention and support to elementary science instruction. Thus, eliminating the materials center would be tantamount to cutting the program. As a result, one can speculate that other areas of the program that are equally important but less visible and concrete (e.g., professional development) are targets instead.

It is obvious that instructional materials themselves are a core part of any hands-on science program. Less obvious are the impacts that the processes for selecting, distributing, and managing them have on other aspects of the program. While providing teachers with the most concrete illustration of science instruction, from a broad outlook, the issue of instructional materials links to many of the other factors influencing sustainability, including professional development, leadership, money, and perception. Ultimately, together they influence the programs' sustainability.

LEADERSHIP

- *The requirements of a sustained program's leadership vary at different stages of the program and with shifting district conditions.*
- *The style of leadership needs to coincide with the culture of the community and the needs of the program.*
- *Attempts to develop the engagement of school-level leaders have largely been unsuccessful.*

- *Superintendents have three tools they can choose to exercise or not: authority, political influence, and budgetary influence.*
- *Program leaders and their leadership teams are ambivalent about the more supervisory and coaching roles they might play.*

Leadership in sustained programs is wide ranging and evident at all levels of the system. It extends 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). Leaders of the nine programs in this study have had widely varied strengths and weaknesses, but their ultimate success has been dependent on their abilities to be flexible, respond to shifting district conditions, and interact appropriately with the local culture. Their experiences have offered insights into how leaders at all levels in a district can contribute to sustained programs.

The most prominent leader in each of the sites has been the program leader. Most often a K–12 or elementary science coordinator, these program leaders have shared several important characteristics. First, they all have been passionately devoted to the vision of a hands-on science program. These individuals have not viewed the job of science coordinator as a stepping stone to more expansive work, although some eventually moved into assistant superintendent positions. Rather, they have been committed to hands-on science, and once they took on the job of leading a hands-on science program, they remained, many for several decades. While the program leaders have had widely varied backgrounds (of the nine original leaders, three did not have science training, and three others did not have significant elementary classroom experience), they all have been intelligent and passionate about their work, with the management skills to enable them to realize their visions, albeit with different styles and approaches.

As the discussion of culture suggests, these programs were initiated in districts with unique cultures, and each program leader relied on his or her ability to reach out and respond to that culture. In Sycamore and Hudson, for example, the districts are very cohesive, and colleagues work closely together, relying on the trust they have built over time. Thus, in Sycamore, Stowe generated the right opportunities and context for the program to move ahead by leading more by quiet example than by charisma and exhortation. In Hudson, Lawson has been a wellspring of energy and determination. Her guiding principles, such as, "It's amazing what you can do if you don't care who gets the credit," and "Always make sure that other people profit significantly from what you do—but that doesn't necessarily mean financially," have led her to make decisions that have brought favorable outcomes to her colleagues as well as to her program.

In Benton, on the other hand, the small, urban district with a mix of high-profile universities and high-tech corporations is characterized by its decentralized organizational structure. In this context, Constance Connor, with her straightforward and forceful personality, was seen as a "visionary" and a "powerhouse," who was able to advance the program at all levels of the system. In Garden City, another decentralized system that was larger than Benton and had a much more bureaucratic structure, Fran Reece was able to bypass many of the structural obstacles and develop collegial relationships with teachers, principals, and others. She was seen as one of the most accessible and responsive curriculum consultants in the district, which was highly valued. Stowe, Lawson, Connor, and Reece all reflected different strategies and approaches that were well suited and successful within the context.

In maneuvering throughout their district's systems, program leaders have had to garner support at the central office level, the classroom level, and every level in between. Credibility, as both an educator and a content expert, is important in this regard. Program leaders have established their credibility in different ways, either through their own credentials, by building on them as they went, or by attracting other experienced leaders to the program. When Evers came to Bayview as a high school chemistry teacher to lead the elementary program, he spent the first several months in elementary classrooms, sometimes substituting for extended periods. In Benton, Connor was also missing elementary experience, but she brought talented classroom teachers on board as resource teachers to meet that need. Pearl North, on the other hand, also without elementary experience, gathered support for and interest in the kit-based program through her innate ability to draw people in and get them excited about the materials.

One notable leader in this regard is Dorothy Parson, who came to Bolton as a reading teacher with minimal science training. As she gradually grew into her position, she gained knowledge, but when it came time to redesign the curriculum, she gathered a core team of teachers to undertake this process together. This strategy built credibility, not just for her in her role as program leader, but also for the curriculum itself and for the professional development that would soon follow.

As these leaders developed the programs in their districts, they also attended to their own professional development needs. Among them is an interest in participating in a support network that reaches beyond their districts to leaders of hands-on science programs in other places. All of the district leaders spoke of the importance of connecting to others who are engaged in similar work, and through those relationships they have found stimulation, inspiration, and lasting sources of advice. The early pioneers found each other and built their networks piece by piece, while leaders of the younger programs spoke of the importance of the NSRC[10] and NEXT

---

[10] NEXT STEPS was originally sponsored by the Association of Science Materials Centers (ASMC) and now is jointly run by ASMC and the National Science Resources Center (NSRC).

STEPS conferences. In any case, all leaders spoke of the sense of isolation they often feel and of the importance of having the sustenance that colleagues and mentors provide.

Different leadership skills are required for the various stages of program development: establishment, maturation, and evolution. In other words, program start-up, maintenance, adaptation, and growth each calls for leaders to capitalize on particular strengths and abilities. Establishing a program requires the ability to sell the idea at the various levels of a district, as well as the ability to create the structures—professional development and materials management systems, for example—that enable the program to function.

Although the tasks remain consistent, each district's culture and operating systems require different strategies to accomplish these tasks. For example, in Montview, Donahue worked methodically to develop understanding of and comfort with a hands-on approach to teaching science at all levels of the district's structure. He developed his ideas in writing and included them in district materials, thus moving kit-based science in Montview to standard daily practice. Glenwood also faced the challenge of establishing a common understanding of and commitment to kit-based science. But there, leaders required skills completely different from Donahue's. Glenwood leaders needed to have the ability to organize many disparate players—the varying grants—and then move them all in the same direction.

Moving a program past the establishment phase and through the maturation phase requires yet a different type of attention. A defining characteristic of successful maturation phase leaders is how well they manage administrators and district turbulence. Since the program and the context change require different strategies, leaders have to be flexible and have a variety of strategies and tools. Two bodies of leadership skills are necessary: the ability to make programmatic decisions and the ability to make political decisions. The shift from one body of skills to the other can sometimes be accomplished by one leader, but occasionally it can be more readily attained when a change in leadership takes place.

For example, in Bolton, Pearl North had the political savvy, energy, and charisma to begin a program where there was none before. She was persuasive and expansive, creating the interest and commitment in the administration to support something new and complex. However, it was difficult to sustain the momentum of the establishment phase, and the program eventually waned. Dorothy Parson's vision, deep understanding of classroom instruction and professional development and ability to manage the details of a large grant, enabled her to reinvent the science curriculum and reinvigorate the program as well. In Garden City, Reece observed that particular political skills were necessary to meet the needs of her program as it moved past the 10-year mark. With a new superintendent, the impact of high-stakes testing, and her questions about how the program should continue to develop and strengthen, she recognized the importance of being "astute." Taking the program to the next level in Garden City would require considerable political savvy and the ability to work through the system.

Program leaders also recognized the importance of engaging school-level leaders in the science program. Their programs include plans to engage principals and cultivate their interest in and willingness to advance the science program in their schools. However, principals are more often regarded as administrators rather than educational leaders, and science is not a high priority, particularly in light of recent testing in math and language arts. As a result, these attempts are usually met with interest but they are difficult to sustain and are largely ineffective over the long haul.

In Lakeville, for example, Wolters had included monthly meetings with principals as part of her strategy to keep them informed about the science program, give them professional development, and assist them in supporting science within their schools. Over time, however, these meetings were reduced to four a year, then to new principals only, and finally Wolters' access to principals was completely removed. It is worth noting that part of the reason Wolters had access at the outset was because of the energetic support of the assistant superintendent for curriculum and instruction. Principals were responsive to his advocacy for the program, but when he left the district and another champion for the program at that level did not emerge, principal involvement dwindled.

Most programs also have devised strategies to provide instructional support to teachers at the school level; however, these too fall short of their goal. Teachers in leadership roles typically receive additional training and professional development to enable them to seek out colleagues who need assistance and provide them with advice and guidance. In Benton, where the decentralized structure makes school-based support very important, each school has a "liaison" between the science department and the teachers. They offer support to classroom teachers, help colleagues who are very resistant to change, help teachers find materials distribution information, and also help manage materials. In Glenwood, "lead teachers" receive four release days per year to facilitate science teams on site. They present units to their teacher colleagues with the help of a partner scientist and explain district science assessment programs to teachers, the principal, and parents. Some also teach during summer kit training sessions.

Regardless of the investments in professional development, teachers ultimately feel uncomfortable in school-level leadership roles. Bridging the gap between colleague and coach is difficult for them, and they eventually gravitate toward clerical and logistical tasks. The exceptions to this are those schools with previously existing strong cultures of collegiality and a tradition of peer coaching, examples of which are found in Benton and Sycamore. Still, as the discussion of philosophy (p. 83) illustrates, the professional development these leaders received did reap important long-term dividends for the programs' sustainability. First, as mentioned above, these teachers often took on the job of providing kit training outside of their school buildings, an environment in which many of them felt more capable of coaching their peers. Second, with each investment in their understanding of hands-on and inquiry science, they became more committed to the philosophical foundation of the program.

Yet another strong leadership influence rests with the superintendent. Superintendents' support can propel a program to the front burner, while their neglect can relegate it to the back shelf or worse. They influence the program's stature by exercising their power over the budget, accountability measures, and political relationships (e.g., the school board, their assistant superintendents, others in the administration, and the district's principals). Superintendents in these districts took one of three general positions in relation to the science program: (a) strong and supportive, (b) benign neglect, or (c) obstructive.

Superintendents' abilities to influence the well-being of the programs compel leaders to court their interest and commitment in a variety of ways. In some cases, support at the district level is a function of administrators having become familiar with the program years before as a teacher or principal. This is true in Montview, where the deputy superintendent, who provides valuable support, became a fan of the program when he saw it at work in his school as a principal. In the early days of Garden City's program, the leader took the then-deputy superintendent on a trip to visit Mesa, Ariz., and Schaumberg, Ill., to see their programs in anticipation of building a similar one at home. In retrospect, the science coordinator saw that invitation as "a stroke of genius, the best thing we did. Having him on the team was key." He has since become superintendent and a strong supporter of the program.

The pressure to establish this relationship as superintendents come and go can be a considerable drain on the leader. In Garden City, the superintendent left after giving his consistent support at the district level, and the new one has had an entirely different style and perspective. One partner in Garden City observed, "If [the superintendent] decides that science is a priority, the science program will grow. He can't cut the program because the local corporations support it, but he could decide that it's in good hands and disregard it," which would, in effect, cause it to stagnate.

In addition to reaching out to the central office, program leaders also have built "mid-level" leadership structures to increase the capacity of their programs. Most important, leaders need help in making strategic decisions that will guide their programs forward, and that means adding to their ability to stay informed about the programs' status, problems, strengths, and weaknesses. The limitations of accomplishing this feat have already been discussed, but still, leaders have devised a variety of roles and structures to address this need. In Glenwood, there is a team of resource teachers who guide the elementary program together; in Benton, resource teachers provide the direct outreach to schools and also work with Connor; and in Hudson, resource teachers work with the leader and are informed by three other types of positions that provide direct support at the school and district level. In almost all cases, these additional leaders are former classroom teachers, which present some distinct positive and negative implications for the programs.

As mentioned earlier, teachers have difficulty assuming the role of coach among their colleagues at the school level, and this discomfort is seen with resource teachers and also, occasionally, with program leaders. Teacher leaders' experience as classroom teachers and their familiarity with the pressures and difficulties teachers face have given them credibility with their colleagues, established a level of trust, and also provided a unique entree into their experience using kit materials. On the other hand, it has made assuming a role that suggests added authority very challenging and, sometimes, uncomfortable. District- and school-level teacher leaders were ambivalent about the more supervisory and coaching roles they might play, which explains at least in part, why teachers found it easier to do off-site kit trainings than in-school mentoring.

Therefore, although leaders expressed some frustration with the lack of formal authority they have within the district, it has afforded them the opportunity to assume the role in which they are most comfortable: helpmate rather than manager. This tension likely exists for all subject areas, but science may be especially affected because of the high need for professional development in relation to kit use and the critical benefits that teachers reap when they share their experiences with kits.

MONEY

- *Supporting a science program with district funds requires vigilance and creativity on the part of program leaders, and commitment from the district's administration.*
- *External funds can boost a program while, at the same time, accentuating existing or establishing new potential inhibitors to that program's sustainability.*
- *Uses of external funds often reflect the interests of the funder and, thus, can influence the shape of the program.*
- *District funds and external sources of support each are associated with particular advantages and challenges that need to be accounted for within the context of the district's culture.*

Many equate program sustainability with a district financial commitment. While there is no question that money is a critical player in a sustained program, its role is far more complex than the simple presence or absence of financial resources. The source of the money, the amount needed, the way it is used at different points in the developing life of the program, and finally, the nature of district culture and interactions with regard to money all are significant issues.

Funding for each of the science programs in this study has been a complex amalgam of resources, including Eisenhower funds, donations from partners, money earmarked for textbooks, external grants, and general district fund line items. Identifying and tracking the varying sources of funds was a challenge, even for some of the program leaders, indicating that the business of securing funds for a program, even when restricted to within-district resources, is a complex job that requires attention and creativity.

Some leaders, such as John Evers, were well-positioned to oversee and manage the various sources of money. From his earliest days in Bayview, Evers made it a point to pay close attention to the budget, and not just the science program's but the district's as a whole. In this way, he was able to maintain a close understanding of how the budget was structured, how the various curricular areas fluctuated in support over time, and how their budgets related to their educational goals and needs. With this knowledge, he was equipped to be an effective advocate for science because he was able to consider the science program as a whole within the district. Linda Lawson, in Hudson, also was in a similarly strong position. There, the program was funded in conjunction with health, social studies, traffic safety and world languages. Thus, as the program leader with oversight over all of these areas, Linda had the authority to move money from one place to another as needed. Aided by the trust that she built among the district's administration over the many years they worked together, Lawson had the respect and freedom to expand and contract her science budget as program needs and district resources allowed.

Other program leaders have been less fortunate, having to seek more creative ways of rallying financial program support. Fran Reece, for example, operates in a large organizational structure where she has no direct access to district professional development funds nor does she have sole oversight of Eisenhower money. A further challenge is that the Copper Beech Science Center, an organization affiliated with but independent of the district, runs its own professional development and not only controls its own budget but also, occasionally, competes with the district for funds. In spite of these challenges, Reece has been able to capitalize on her relationships with administrators and her excellent communication skills to ensure that the Garden City Board annually approves district budget line items for kit refurbishment and materials center management. Constance Connor also has had little direct oversight over district budget decisions. During the years of her NSF grant, she was able to design and implement professional development somewhat independently. Once those funds were expended, however, Connor risked the loss of her science resource teachers. Each year, Connor reshaped and redefined the resource teachers' positions so that they were better aligned with district priorities. She has had modest success in retaining her staff, though the recommitment of district funds to support them remains in question.

Most programs in the study established themselves primarily with internal district funds (Hudson, Montview, Bayview, Garden City, Sycamore, and Bolton). Several of these sites have supplemented their programs with relatively small grants, and several have sought and received larger NSF grants at later points in their development. Benton, Glenwood, and Lakeville, on the other hand, have had a relatively large amount of external support concurrent with their efforts to establish their programs on a large scale (Benton had a large NSF grant; Glenwood had support from a local nonprofit organization, and not long after, benefited from NSF funds; and Lakeville began with several small grants followed up with an NSF grant). This range of funding strategies suggests that no single approach is more effective than another in ensuring a program's sustainability. Rather, it highlights some of the advantages and obstacles associated with both a reliance on internal funds alone, as well as the acquisition of external grants.

Regardless of the developmental phase in which a large grant is secured (establishment, maturation, or evolution), the influx of money enables districts to accomplish large tasks in a relatively short amount of time. In Bolton, for example, the LSC grant strengthened their revamped program by supporting professional development that introduced all 1,080 teachers in the district to the new kits and to the philosophy of hands-on, inquiry-based teaching for four years in a row. In Glenwood, where external grants are a consistent source of support, an early mix of projects allowed the district to train over 100 teachers in classroom implementation of hands-on science kits, develop a cadre of nearly 30 teacher leaders for school-level support, and raise principals' awareness about science instruction. And in Lakeville, after the program had been developed, field-tested, and expanded to a small degree through small grants from a variety of sources, a large NSF grant enabled the program to be disseminated to all schools in the district at a much faster pace.

In addition to the financial benefits of grants that have supported these major undertakings, external funds also bring additional independence, stature, and influence to the program leaders. However, they can exact an internal price. Money is not often plentiful in school districts, and when one program comes into a large sum, it automatically creates "haves" and "have nots." In Bolton, for example, where communication between departments does not always happen with ease, the arrival of funds created subtle tensions that exacerbated communication problems. Even in Glenwood, where multiple grants have been directed within the area of science, rivalries have emerged between the leaders of each of the initiatives, requiring careful intensive attention to improving their collaboration.

Even as the grants bring opportunities to the programs, the program leaders have to address some challenges associated with the changing ebb and flow of funds. Though most districts feel they never have enough money to support their programs, those with large grants have more to spend than others and, thus, perceive the end of those resources as being a loss for the program. For example, compared with Sycamore or Montview, Benton and Lakeville have had great wealth. Still, when their external funds were expended, momentum was lost, leaving program leaders feeling like they were taking steps backwards rather than merely slowing the pace of growth or even holding their ground. In Benton, the loss of funds was particularly acute as it coincided with the arrival of a new superintendent and the advent of a high-stakes test. Together, these three events generated tremendous pressure on the program, creating significant tension and uncertainty about its future.

Another constraint that accompanies the benefits of external funds is the accommodation of funders' guidelines. Generally speaking, funders may wish to focus on particular grade levels, specific sets of materials, or target sub-populations of students or teachers. As a result, they can be prescriptive regarding acceptable curricula, recruitment for participation, and professional development design. The LSC program, for example, required a minimum number of hours of professional development for each teacher supported by the funding. These funder requirements may or may not meet the original interests or needs of a district's program at a particular time but compel the leaders to adapt accordingly.

Glenwood offers a case in point. This district program has enjoyed the support of several grants since its inception. In fact, seeking external funding seems to have become the norm for the science program. Sondra Calder recounted her meeting with a newly arrived superintendent in 1989, explaining, "He basically told me to do whatever I wanted, that I was the expert. He'd support me with anything except money." His verbal support was, indeed, critical, and Sondra went on to capitalize on it with successful grant writing. Still, while the positive outcomes of its grants were obvious, Glenwood faced some resulting challenges. First, Glenwood seemed to have developed some dependency on external funding and faced uncertainty about where their next dollars would come from. Such uncertainty has not even occurred to leaders in programs such as Sycamore and Garden City, where the funds have been embedded in the general district budget. Furthermore, with the arrival of each grant, the science programs have had to make an adjustment to accommodate the funders' interests. Thus, the Glenwood program has had to periodically re-define its strategy and goals. While this, perhaps, has disrupted the consistency of some strategies, Glenwood has held fast to its fundamental core belief in teaching science using a hands-on science program.

Districts that avoided the problems of seeking and receiving external funds have taken pride in their self reliance. Sycamore, Garden City, Hudson, Bayview, and Bolton have been able to rely on district funds alone for many years because they and their districts are deeply committed to the science programs. Although funds have certainly fluctuated in all of these places, the science programs are accepted practice and, thus, receive consistent support.

What leaders gain in avoiding the pitfalls of external funding, however, they lose in the ability to make large-scale impacts on their programs in short periods of time. In Hudson, this constraint has been a good match for the style of the program leader. Lawson has been inclined to "soft sell" the program to teachers and engage their interest gradually rather than take them by storm. However, because her budget has been so limited, she also has had no other choice except to grow the program incrementally over time. The same might be said of Garden City, where the leader also has been inclined to a "soft sell" approach. However, in this more insular district, self-reliance is a blessing and a curse. It has provided the program with steady support since its inception, but it also has kept the leader isolated from the larger science education community.

The importance of steady financial support for the sustainability of hands-on science programs is the only simple conclusion that can be drawn from these nine districts. Funding had been secured in a variety of ways and from different sources, with each funding strategy accompanied by implications for the program and the program leader's role within the district culture. These sites suggest that there is no single way or best way to fund a hands-on science program that will ensure its sustainability. Rather, it is the leaders' abilities to understand and address the complex nature of securing financial support that is key.

PARTNERSHIPS

- *Typical partnerships are somewhat superficial and supplemental but still serve to enrich the science program.*
- *Deep partnerships are rare, require investments of resources and political currency, and can have both positive and negative impacts on the sustained science program.*

Districts in this study have had partnerships that fall into two broad categories. Most common have been the "limited" partnerships forged between a local business or organization and a single school or district area. All districts have had many partnerships of this kind: Garden City has 1,500 at last count. With specific regard to science, these partnerships most often focus on supplementing the core science program, including providing resources for purchasing additional science materials, providing volunteer help, and making available space and materials for school meetings. Albeit somewhat superficial, these partnerships still enrich the programs, and, particularly for schools in high poverty areas, often provide much-needed assistance. Though ubiquitous in the districts in this study, these partnerships do not seem to have far reaching implications for contributing to or inhibiting sustainability.

The other category of partnerships encompasses those that have been deep and comprehensive. Such partnerships are rare, occurring mostly at the district level and requiring investments of resources and political currency, as well as shared planning and leadership. Partnerships that fall into this second category offer valuable insights into how collaborations that emerge through partnerships play significant roles in sustained programs. Among the nine districts, Lakeville, Hudson, Glenwood, and Garden City each have had a part in a partnership that was vital to the program's establishment and/or maturation. Although each partnership has functioned differently, they have several features in common, chief among them their high cost in investments of time, attention, resources, and political capital. Although partnerships of this kind are generally lauded as assets without reserve, data collected in this study suggest that while benefits are clear, partnerships also add complexity to reform efforts that can slow or even divert a program's momentum.

The program in Lakeville, for example, began as an invention of two professor-scientists at Grossen University who wanted to introduce hands-on science to the district. They brought their idea to the superintendent and district science coordinator, and in 1986, formed a partnership with a pilot school. Working with the school's improvement coordinator, they developed and field-tested a kit-based science program. As the program grew, the professors and their organization, Grossen University Science Outreach (GUSO), provided valuable guidance and professional development, helped the district secure NSF funds to support the program's initial expansion, and provided ongoing help with finding funds to support resource teacher positions. GUSO grew along with the program, and with its own NSF grant, replicated the model it had developed in Lakeville in other districts in the state.

The benefits of this partnership are obvious, but there also were unexpected challenges that drew time and energy away from the program's progress. First, it was not easy to establish the partnership because, as one of the professors put it, "We had to understand the cultural differences between the confrontational world of scientists and the nurturing world of education." With diligent effort from the professors, an assistant superintendent, and the pilot school improvement coordinator (who ultimately became the science coordinator for the district), they eventually became an authentic team.

Second, the role of GUSO with regard to funding for the science program was difficult for the central office administrators to decipher. To many in the central office, it appeared that the program was supported by GUSO, which suggested that increasing district support beyond its current level was unnecessary. GUSO had gone out of its way to find ways to fill the gap in district funding and took the administration's unwillingness to increase the budget for the science program as a lack of commitment to science. Ongoing tensions on this topic of who should and could support the science program sowed the seeds of resentment and mistrust between the parties.

Third, GUSO grew as an organization and gained a national reputation, in part, by expanding on its early work in Lakeville. This caused further resentment in the district and aggravated the existing ill will. The cumulative effect of the strained relationships ultimately caused GUSO to withdraw from its direct work in the district, although the professors were still informally available to the science coordinator. When controversy arose about the value of the science program in 2000 and erupted in a vigorous public debate, the role of GUSO was a target of some of the resentment that was expressed. As public support was successfully rallied for the program, GUSO stayed in the background.

In Garden City, the science program has engaged with two major partners. Some aspects of the partnerships have offered significant benefits while others have been more troubling and draining on the program resources. The first organization, Copper Beech Nature Center, is an extensive and elaborate facility. Copper Beech receives some of its funds from the district, and its mission is to serve as a resource for the district's teachers and students. Its professional development workshops are popular because they provide a wealth of materials and activities that teachers can bring back to their classrooms. Conflicts have arisen, however, because the Center's educational staff has been reluctant to coordinate the professional development agenda of Copper Beech with the goals of the district's program, which, in turn, has created confusion for teachers and undermined some of Reece's work in the district.

It also has created frustration as Reece has lost the opportunity to see a much-needed professional development program made available to Garden City's teachers. As a final source of conflict, Copper Beech and the district have sometimes competed directly with each other for funding from the same sources. As a result, where they could have been constructive partners and coordinated their efforts to produce an outcome that neither alone could have achieved, they have been working at cross-purposes despite intermittent efforts to break the cycle. As Reece said, she wished there were a way to "force them together," because the continued competition and lack of connection could drain energy and enthusiasm and deprive both of numerous opportunities.

Garden City also has a partnership of a different nature with SecCorp, a multinational corporation that came to the community in 1992. Upon their arrival, SecCorp leaders sought opportunities to play a role in state and local science education for two main reasons. First, SecCorp's work was linked to the field of science and, thus, science was a substantive area of interest. Second, SecCorp was interested in capitalizing on the public relations value a partnership with the school district would provide. Over several years, Reece had built a sound relationship with the corporation that gleaned several kinds of rewards. Most obvious have been the financial contributions SecCorp has made to the district's science program. But, in addition to funds, the partnership also has provided Reece with other benefits, including facilitating her involvement with several statewide planning institutes and assisting her with the adoption process by hosting an introductory reception for K–8 principals and by meeting with the new superintendent to share SecCorp's endorsement of the science program. Reece noted that the partnership with SecCorp has offered tremendous benefits of support and credibility, but said that these benefits have come at a considerable price. In a district that is so insular, going "outside" to develop partnerships has been an issue that has had to be carefully negotiated. For this reason and others, time devoted to nurturing the relationship has been substantial and has drawn away from other work of immediate importance to the program.

Glenwood's science program has fostered many relationships over the years, but two are notable for their extensive involvement during the program's early years: the state university and the science museum. At the local state university, the Science and Health Education Partnership (SHEP) was initiated in 1987 by a professor of biochemistry and biophysics. SHEP's mission was to improve the quality of science instruction in Glenwood, and its principal avenue was professional development for teachers through building an integrated community of scientists and educators. SHEP had 12 different programs utilizing approximately 350 volunteers from the university, including students, staff researchers, faculty, and post-docs. These volunteers provided about 10,000 hours of service per year to about 75 percent of the elementary schools in Glenwood, where they worked closely with teachers to improve their science instruction.

The science museum has long been a prominent feature in Glenwood's cultural and educational life. Known nationally as a center for inquiry-based science education, it offers professional development to 500 of the district's elementary and secondary teachers. Teachers in Glenwood have long relied on the museum as a key resource for science; however, a formal partnership between the museum, SHEP, and the district didn't begin until 1989. At that time, SHEP and the museum coordinated a series of summer institutes for 27 elementary teachers intended to form a cadre of teacher leaders.

At the beginning, tensions arose over turf, budgets, control, differing agendas, and different cultures. Eventually, in an effort to create some cohesion for the district out of these disparate projects and partners, the superintendent asked Sondra Calder, the K–8 science coordinator, to draw the players and the projects together. From 1991–93, she facilitated regular meetings between the district, museum, and SHEP. They became known as the Science Council and their goal was to create a vision for how the three projects could link and build onto each other.

The partnership improved, but it took some time and attention before each organization set aside its own agenda and was able to accept common goals, appreciate the unique assets that each partner brought to the work, and work within the confines of their different cultures. A capstone to this hard work was achieved in 1994, when SHEP and Glenwood submitted a successful proposal for an LSC grant. It became clear to the director of SHEP, who was the principal investigator, that the district was the more appropriate lead. It took a year of negotiation with NSF, but ultimately SHEP transferred the project to the district, and SHEP became a subcontractor on this $2.5 million grant. Even though the shift made sense given the nature of the work, it also meant a loss of funds for SHEP. This difficult decision was truly appreciated by the district and was evidence to all that the collaboration had finally become an authentic partnership

The last district to provide an example of a partnership is Hudson, where the relationship with the local state university was informal for many years. Lawson, the district science coordinator, knew a professor there personally, and they conferred on a casual basis. A more formal partnership formed in 1996, when Lawson wanted to apply for NSF funds. She approached an astronomer and university professor to inquire about her interest in sharing responsibilities as principal investigator on a local systemic initiative grant. The professor agreed and, as she said, "The day the money arrived is the day the partnership began."

This partnership is markedly different from Lakeville, Garden City, and Glenwood in that it is minimal, limited to opening up space for some Hudson teachers to enroll in the professor's inquiry-based science course and lab during their pre-service program. This partnership requires little of the investments that were problematic in the previous examples, but there were minimal benefits for the district as well. Still, this type of partnership represents a very common scenario between a district and an institution of higher education, and for that reason, deserves acknowledgement here.

These different experiences share common lessons regarding the role that partnerships can play in the sustainability of a science program. While there is no guarantee that a partnership will result in adding to the sustainability of a program, the partnerships that do offer significant contributions require investments of time and hard work. An associate superintendent in Glenwood spent about 20 percent of her time on building and maintaining partnerships in the district, principally bridging gaps between cultures of schools and businesses, universities, and other organizations. Moreover, these experiences suggest that even with diligent effort, the risks are great for political conflicts to arise that are beyond the capacity of a district to smooth over. As with many of the other factors found to be significant to sustainability, partnerships are a component that can have positive and negative effects, depending on the context and conditions in a district.

PROFESSIONAL DEVELOPMENT

- *The roles of specific approaches to professional development in sustained programs vary, depending on where the programs are in their evolution.*
- *Professional development needs perceived by program leaders are not necessarily congruent with the needs perceived by teachers, nor are they necessarily the activities that will support the sustainability of the program most effectively.*

- *Professional development contributes to sustained programs independent of its impact on classroom practice.*
- *Teachers trained to provide professional development support at either the school or district level often represent unrealized potential.*

Professional development in the context of hands-on elementary science programs refers to activities focused on increasing teacher, principal, and administrator capacity to understand and implement hands-on, inquiry-based science in classroom or school, grasp the scientific content of particular units or lessons, and manage materials and student interactions with those materials. Such activities might include mandatory or voluntary trainings on kit use, summer academies focusing on inquiry teaching methods and/or science content, study groups entailing individual exploration of science questions or student learning, and follow-up debriefings on kit use in the classroom. In the absence of clear data on the impact of specific professional development activities on classroom practice or student outcomes, this study explored several other avenues for understanding the role of professional development in sustained hands-on elementary science education programs.

Given the range of possible professional development strategies, program leaders need to make decisions about which strategies to choose, and when to use them. One strategy employed by all districts has been basic training on how to use the kits. In six of the nine districts studied, this professional development has been mandatory but at different points in the program's evolution. For example, in Lakeville, kit training has been required from the start. In Bolton, however, mandatory kit training didn't take place until the second generation of the program, nearly 15 years after the program was originally established. Similarly, when Hudson received NSF funds, they initiated mandatory kit training for the first time in 20 years. As demonstrated in this variation, while it is clear that kit training sessions are essential, all teachers need not be required to participate in them to establish a program.

Still, even in those sites where kit training was not mandatory, it composed the bulk of the professional development in the establishment phase. And, in Bolton and Hudson, although widespread kit training arrived long after program establishment, it coincided with the entry of that program into a new stage of development—re-establishment. Beyond initial establishment and these occasions of program re-establishment, kit training has not maintained its presence in the sustained programs. In most cases, including Benton, Garden City, and Bayview, kit training has been available but voluntary and, for the most part, poorly attended. Still, these programs continue to endure, which suggests that while kit training is important at establishment, it seemed less so for program growth and evolution. This may be a result of an already well-established base of participation, an acceptance of the program as standard practice, or the fact that other mechanisms function to fulfill the same purpose as kit training. It also is worth noting that none of the sustained programs approached establishment of their reforms by first focusing on the philosophy and pedagogy of inquiry and then turning to kit implementation. Rather, all focused on simply getting the program in place and reserved their efforts to improve quality for later.

Sycamore, for example, has very much de-emphasized kit training except for periods when the district introduces a new kit, in which case training on that kit is mandatory for all teachers, with optional follow-ups. However, in the opposite case—when a kit is established but the teacher is new—that teacher does not have the same opportunity for kit training. Rather, new teachers are expected to rely on their peer coaches and grade-level colleagues for support. This approach reflects the fact that until very recently, Sycamore had a notably low rate of teacher turnover, rendering one-on-one attention to new teachers feasible.

Indeed, teacher turnover is a mitigating factor in all professional development planning. Those places with high levels of teacher turnover have been inclined to focus relatively greater amounts of resources on providing induction-level support. Teacher turnover also affects the potential return on higher levels of professional development. Summer academies and study groups, for example, require large investments for relatively small numbers of teachers. Participants often are the teachers who, in turn, assist in kit trainings and provide support for newer teachers. When teacher turnover reaches these teachers—whether leaving for better paying jobs or reaching retirement age—the program suffers a loss, not only of those more experienced teachers but also of a pool of potential professional development leaders.

Another important issue is the level of perceived need for professional development among teachers, administrators, and program leaders. In some cases, teachers' and program leaders' views of professional development needs are well-matched, though not always with the most productive outcomes. For example, Fran Reece expressed uncertainty about the extent to which the program needs to emphasize "inquiry" over "hands-on" and her resulting ambivalence about communicating more directive goals. As a result, the professional development program there remain a simple base of voluntary kit training with other options for teachers in the form of one-time workshops offered by individual teachers through the district, the SSI, or the Copper Beech Science Center. Her uncertainty is mirrored, to some extent, by teachers' lack of interest in what other districts referred to as "advanced" professional development. Across the district, teachers expressed widespread satisfaction with the program as it is. Teachers who were informally identified as leaders are confident in their abilities to teach the units and remarked that the small-scale, voluntary kit training is sufficient. They made no suggestions about possible areas of improvement, such as increasing content knowledge, looking closely at student assessments, or reflecting on their instructional practice, which raised in other districts.

Hudson, on the other hand, stands in stark contrast to Garden City. There, almost everyone interviewed expressed a desire for increased professional development for teachers and principals. Hudson offers a wide range of professional development for teachers beyond kit training, which is required for new or new-to-grade teachers. New teachers also are encouraged to participate in a program built on a four-hour session where they discuss the science kits with what Hudson calls "liaison" or "mentor" teachers. Additionally, Hudson offers workshops in various content areas and provides site-based, in-service training. Some of these offerings are shaped by the LSC, but nonetheless, reflect the interests and perceived needs of the program on the part of the program leadership.

Sycamore illustrates a different issue related to alignment of perceived professional development need. Science coordinator Allison Stowe has provided a range of upper-level, voluntary professional development activities targeting eager teachers and identified teacher leaders. A major focus has been summer inquiry institutes, with four follow-up meetings during the school year. Stowe also has been working with a local university to create new credit courses for both in-service and pre-service teachers. However, while teachers are interested and eager, they aren't necessarily intellectually focused on the goals and outcomes of the professional development as Stowe and her leaders see them. It seems that there is a conflict for Stowe and other leaders as their programs move from establishment to maturation to evolution. They expect and push for more development (e.g., pedagogical skills) in their teachers' understandings, and yet their goals may simply be more than the teachers care to (or need to) pursue.

One would presume that a lack of congruence between professional development offerings and teachers' perceived needs would lead to wasted professional development resources. And yet, in spite of less-than-ideal professional development participation and engagement, programs endure. A partial explanation may come from the notion that professional development can contribute to sustained programs independent of its impact on classroom practice. Program leaders of districts in this study uniformly have no systematic means to assess the impact of particular professional development strategies on classroom practice. But the return on the investment of those practices, particularly those considered "upper-level," expand beyond the obvious intent to increase quality of classroom practice. Advanced professional development has a positive effect on individual teachers in terms of deepening their commitment to the program and strengthening their philosophical understandings.

In Benton, for example, district-level resource teachers and school-level liaisons all recognize and express their appreciation for the high-quality professional development support they receive. In fact, the science department's professional development has been mentioned as a model for other subject areas in the district. The training is much appreciated, and it cultivates in these participants a group of loyal program supporters and, perhaps more importantly, true believers in the program philosophy. As a result, these teachers have become formal and informal human resources, disseminating support for and belief in the program philosophy. Based on this study's definition of sustainability, maintaining the core beliefs and values is key. Thus, the ability of professional development to support those values—even if it has no impact on the quality of instruction—contributes to helping the sustained program endure.

At the opposite end of the spectrum of unintentional benefits described above is the regrettable situation of unrealized potential in school- and district-level leaders. In the majority of cases, school-based teacher leader strategies have not succeed in promoting mentoring or other collegial support structures. Teachers often feel uncomfortable in that role, sometimes preferring to focus only on logistical functions. As a result, despite training to enable them to give instructional support and guidance for their teacher colleagues, school-level leaders are commonly found doing clerical work. In Bolton, for example, the consensus among all of the program leaders was that the promise of the school-level science consultants is never fully realized. At the end of their LSC project, the school science consultants were primarily assisting with kit orders, a minimal function compared with the broader vision of their potential. The exceptions to this typical situation are particular schools that already have strong cultures of collegiality and a tradition of peer coaching, such as some of the schools in Benton and many schools in Sycamore.

Benton also struggles with similar issues at the district level. While district-level resource teachers have the potential to increase the quantity and quality of science instruction in the schools, cultural and logistical constraints (see more in discussion under leadership, p. 51) prevent them from realizing this potential. They have no real decision-making authority, nor do they necessarily want any. Thus, they are not in a position to act proactively and go where their support is most needed. Instead, they go where invited—one can only presume that they visit the classrooms of teachers actively engaged in the science program. And yet, as explained above, even when professional development investments in these leaders is not fully realized, they still net long-term gains by supporting the growing numbers of individuals who believe in the program.

Hudson seems to be an exception to the constraints on school-level leaders in that there, several levels of support systems (resource teachers, liaison teachers, mentor teachers, and connection teachers) are active. Resource teachers provide in-service training at schools during the year and lead summer professional development institutes; liaison teachers serve as conduits of district information; mentor teachers work with new teachers on curriculum and pedagogy; and connection teachers are responsible for intensive training in science for new and new-to-grade teachers. Resource teachers and school liaisons have been active since the early program; mentor teachers and connection teachers were new with the arrival of the LSC funds. The extent to which these professional development supports will remain intact beyond NSF funding remains to be seen.

In closing, the role that professional development plays in sustainability is somewhat unexpected due, in large part, to the fact that its intended impact on actual classroom practice is unknown. Still, it appears to have an unintended but no less significant relationship to the sustainability of the programs in this study. This is primarily due to its ability to foster deeper understandings of and commitment to the programs' underlying purpose. This was particularly true for teachers who participated in "higher-level" professional development because they immensely appreciated the messages of respect and professionalism that were implied through their participation in those events.

SECTION 3: FACTORS THAT PERTAIN TO THE WHOLE SCIENCE PROGRAM

ADAPTATION

- *No district is static. Thus, science programs must adapt if they are to endure.*
- *Sustained programs are altered in a wide variety of ways for a variety of reasons.*
- *Adaptations can be proactive or reactive.*

The definition of sustainability presented in this study suggests that sustained programs use their core beliefs and values to guide adaptations to change. The earlier discussion of what sustainability is and the phases that programs move through asserts that programs must move beyond establishment and maturation of a particular design to a state of evolution in which elements of the program can vary greatly from the program as originally conceived. It is in this movement—from maturation to evolution and beyond—that programs demonstrate the flexibility and resilience essential to their survival in the ever-changing and, sometimes, volatile district environment. Indeed, every program in this study underwent adaptation.

Some of the most visible adaptations are evident in changes to the instructional materials themselves. In some cases, such as Garden City and Bolton, which started with teacher-developed materials, leaders gradually introduced commercially available units. Other places, such as Sycamore and Montview, which began with materials adopted from other sources (science museums, other school districts, and the early NSF-supported materials from the 1960s), also eventually incorporated newer and commercially available units. A third scenario was found in Benton, which has used commercially available materials from the start, but has added and deleted kits over time and customized some kits to meet local needs and interest. Districts initiate these adaptations for a variety of reasons, including an interest in improving the quality of the instructional materials, a desire to introduce a new concept or topic area to a grade, and more recently, a desire for improved uniformity and greater alignment with state science standards.

Bayview offers a unique perspective on materials adaptation. Until 1999, Bayview had a 30-year history of using exclusively locally developed kits. The first kits were written for the pilot effort in 1966, with a total of 24 units completed by 1968. Because the development of these units was on a relatively short timeline—three years—units underwent development, evaluation, and revision simultaneously, establishing a pattern that continues today. The primary driver of the revision effort was a desire to continually re-examine the materials and improve them. With this long history focused on using locally developed kits, Bayview leaders faced some reluctance on the part of teachers to use the commercially available ones. Still, the leaders feel it was a necessary step because addition of kits is necessary to increase alignment with state standards, and purchasing commercially available kits is the most cost-efficient approach.

Other highly observable adaptations have been in the instructional materials distribution systems. Though primarily logistical in nature, their importance in supporting the endurance of the program should not be underestimated. In the early years of Bolton, for example, the instructional materials distribution system was unreliable, and teachers felt uncertain about whether they would get the kits they ordered. The impact of this uncertainty, one can presume, would be restraint in teachers' commitments to teaching the units and a resulting reduced commitment to the program as a whole. As the program evolved and the materials underwent revision, the system for refurbishment and distribution remained problematic, again creating a potential excuse for teachers wishing to avoid science teaching. It wasn't until 1995, when Dorothy Parson upgraded the kit management process, that the materials center became a stronghold of the program. The new procedures are more reliable and efficient and enable teachers to plan for more precise kit delivery. These changes were concurrent with finalizing the scope and sequence that determined which kits would be taught in which grades and when. Together, these adaptations have helped to solidify the program and embed it more deeply in Bolton's educational program.

Some districts have made less obvious but still concrete adaptations to their curricula, focusing on the instructional sequence. Bolton, Bayview, and Hudson, for example, made a shift from allowing teachers complete autonomy in deciding what to teach and when, to a pre-determined, formal sequence. Specifically, Pearl North's efforts in Bolton focused on introducing the unfamiliar materials to teachers and encouraging their use with gradual growth across the district. Teachers had been able to choose the units they wanted and when they wanted to teach them. When Dorothy Parson stepped in, however, she focused on establishing a more structured curriculum that they would use uniformly across the district. Parson's efforts have been undergirded by her fundamental desire to improve the quality of the program and reflect the district's own views of what is important for students to know and be able to do by the time they finish the elementary grades.

Bayview underwent a similar shift. Along with the introduction of commercially published kits described above, science coordinator Lisa Cooper decided that teachers should no longer select the kits they taught, but instead, should subscribe to a pre-determined schedule for kit instruction. This effort has been driven, in part, by the need to ensure alignment with the state academic standards and the forthcoming state standardized test. Teachers still are able to make special requests for particular kits but are given the clear message that they also need to follow the scheduled units. And finally, Linda Lawson in Hudson made a similar decision, hoping to bring more continuity to the students' learning experiences. Teachers have been willing to give up their autonomy for the shared expectation of what students would learn in each grade.

Another common area of program adaptation is the design and focus of program professional development support. These shifts occurred for a range of reasons (including changing district priorities, leaders' changing views of high-quality professional development, and most often, the arrival of external funds), which illustrate the point that adaptations can be proactive or reactive. More specifically, the resources that come from external grants are substantial and sometimes tied to specific guidelines that reflect the funders' interests. For example, those districts that received NSF Local Systemic Change grants were required to provide a minimum number of professional development hours for each teacher funded for participation. In one scenario, then, program leaders are proactive and seek out funders whose interests support what they hope to accomplish. In another scenario, program leaders desire external support for their programs and reactively adjust their program to align with the funders' goals.

In Glenwood, for example, recent years have brought a shift in professional development from centralized services to a more school-based approach. While the Glenwood program leaders hold a strong vision for the program goals, they have to operate in an entrepreneurial culture that seems to push them into a reactive mode. The LSC grant leaders, for example, hoped to establish a "community of learners" at each school that would enable the LSC program to focus more closely on instruction and make a stronger impact on classroom practice. Then, on the heels of the departure of the LSC, Glenwood's USP program shifted its emphasis to "bridge the achievement gap" and expanded its efforts to include mathematics and science.

Hudson offers another example. Influenced to some extent by the arrival of their LSC, the district has made a shift from more general resource-teacher support to professional development mechanisms available at the school site. Their effort has been focused on "the improvement of science and mathematics instruction through site-based learning communities and cross-district support" and has brought funding for liaison teachers, mentor teachers, and connection teachers. While some funding for liaison teachers existed prior to the grant, the NSF funds allowed that role to expand, and it fully supported the mentor and connection teachers. The extent to which this adaptation in Hudson was a result of the LSC grant guidelines or whether the grant enabled them to go in an already defined direction is not clear. Still, the redefinition of the focus of professional development represents an adaptation in the way professional development was provided.

A final example rests in Bolton, where Dorothy Parson proactively sought out an LSC grant for professional development support to fully implement the newly revised curriculum. She felt that the LSC funds would support large-scale professional development as well as other needs associated with the districtwide curriculum overhaul. Training in the first year of the LSC focused on introducing teachers to the curriculum at their grade level. The second year was intended to focus on more complex issues, such as integration, but the focus needed to be adjusted due to an unanticipated large number of teacher retirements. In the final LSC year, professional development comprised a range of choices, depending on teachers' experience and interest.

As illustrated above, both proactive and reactive adaptations can support sustained programs. Responses to external inputs to a science program might be considered reactive while strategies that are driven by the interests and needs of the project might be considered proactive. All, however, require similar decision-making and action-taking processes on the part of the science leader.

Examples of both proactive and reactive adaptations can be found by looking at districts' responses to the increased emphasis on literacy. In Benton, for example, the science department responded to the district's emphasis on literacy by introducing the use of student science notebooks. The notebooks are intended to supplement and go beyond structured worksheets and encourage students to record and comment on their observations. They are an example of how a reaction to the external pressure of literacy can simultaneously be a proactive step toward improving the nature of the student learning experience.

Glenwood, on the other hand, has taken a proactive approach to the challenge of increased attention to literacy. The teacher leaders of the district spoke of discouraging an "us versus them" mentality and, instead, are constantly looking for ways to connect the science program to the literacy curriculum. Glenwood program leaders note the benefits of focusing on comprehension for both literacy and science. Still, with high stakes resting on student performance in literacy and few consequences for science performance, the tension that comes from fighting for the time in the classroom remains.

Many of the adaptations described above are clearly defined as alterations in choices of instructional materials or professional development designs. However, other less tangible adaptations also guided the evolution of the sustained programs. Program leaders, while still holding to their core beliefs and values, have made adaptations to the program goals, expected outcomes, and their own personal understandings about the extent to which the programs could and should purely reflect inquiry-based instruction. In differentiating between "goals" and "values," it may be helpful to recall the simple definition of core beliefs and values that this study employed, namely, the belief in the value of children actively engaging in the process of doing science.

In Montview, a look over the history of the program reveals ways that the defined goals shifted with national and district priorities at the time. Specifically, at the beginning, the goals of the program were described as "guiding the student in an attempt to find his own answers." Within 10 years, as the program entered a process of revision, program leaders worked to align the program with a newly adopted set of "student outcomes." As a result, the program included more life and human science units and integrated the program with health and environmental education objectives. Within another 10 years, into the 1980s, a district document portrayed the goals slightly differently, stating that "the major outcome of science education…is the development of scientifically literate citizens." They had adapted the specific goals over the years, but the core beliefs and values remained constant.

Bolton offers another example of program goal adaptation that shifted with a change in leaders. The roots of the Bolton program rested with Pearl North and her interactions with NSF-supported programs and individuals who were establishing hands-on science programs in the 1970s. She modeled the Bolton program after their work and focused on getting teachers to teach science and legitimizing the hands-on approach. After the decline of that program and subsequent resurgence, Dorothy Parson took the reins. Her program leadership coincided with Bolton's LSC and targeted improvements in teachers' content knowledge, pedagogical strategies, and teaching skills. As the third generation of the program takes shape with two new leaders, the program goals are likely to shift again. Teacher experts Sophia Harder and Maria Clay are focused on ensuring that the program will continue to evolve and are focused on developing core leaders while ensuring that all teachers understand the "unifying concepts" of the program—emphasizing increased quality of instruction.

Adaptation in program goals and intent are sometimes subtle and evident only in retrospect, even to the leaders themselves. They sometimes emerged only when looking at a collection of program elements over the long-term time horizon of those places that had operated for 20 years or more. Leaders of shorter-lived programs in this study and those just starting out can benefit from the recognition that program goals naturally will evolve and adapt to shifting district conditions and contexts, turnover of leaders, and trends in funding sources. Key, however, is recognizing that throughout all of the adaptations, the core beliefs and values do not waver.

CRITICAL MASS

- *Considering critical mass through the long-term time horizon of sustained programs sheds light on alternative views of what critical mass is and how to achieve it.*
- *In the relative short term, attention to critical mass is highlighted by the challenge of reaching sufficient numbers of teachers.*
- *In the relative long term, attention to critical mass is expanded to include the challenge of obtaining widespread and deep commitment to the core values of the program.*

Discussions of critical mass in reform programs often focus on numbers: numbers of teachers participating, numbers of students reached, and the resource to teacher ratio. This is consistent with a view that one 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. The definition of sustainability generated by this study expands this view to suggest that a program reaches critical mass only when there is a culture of program self-generation. Thus, "critical mass" can encompass other considerations more complex than the simple act of targeting a "magic number" of teachers to implement the program. The data of this study suggest that it also is meaningful to consider critical mass as numbers of teachers and principals who understand and believe in the program's core beliefs and values.

The programs in this study illustrate ways that understandings of and attention to critical mass can shift as a program progresses through each phase of development. In the relatively short-term time horizon of the establishment phase, for example, program leaders of sustained programs have given a good deal of attention to increasing the breadth of their efforts. The most obvious tactic to spread the program to teachers has been professional development focused on kit training—walking teachers through the lessons in the units and helping them understand the science concepts and pedagogical strategies addressed in them. As described in the section on implementation (p. 44), some, such as Benton and Bolton, chose a "snowplow" approach that simultaneously reached out to as many teachers as possible. Others, such as Hudson and Lakeville, chose the "wedge" and targeted smaller groups of teachers in sometimes carefully planned and sometimes looser, more organic approaches. In both cases, the ultimate goal was expanding the program to as many teachers as possible until it became the norm of instruction.

With kit training such an essential part of spreading the word of the science program, teacher turnover has been a thorn in the side of reaching critical mass. In Bayview, for example, of the approximately 600 elementary teachers in the district, 70 of them, or 12 percent, were new hires each year. In Bolton, 10 percent of the total teachers retired in a single year, leaving 450 teachers moved to new grades; while in Glenwood, which will be discussed in more detail further on, 30 percent of the teaching staff were novices in a given year. Theoretically, to achieve the critical mass necessary for sustainability, one would expect that the districts would have a regular, consistent schedule of kit training to bring new and new-to-grade teachers on board. But in most of the districts in the study, this was not the case.

For some, large-scale kit trainings were closely tied to external funds and essentially ended when the funds were expended or redirected to other priorities. For others, such as Garden City, there was little demand on the part of the teachers—even though they had no training, many felt well-equipped to use the materials and, as a result, the workshops that were offered were poorly attended. The situation was similar in Bayview, where program leaders sent personalized invitations to new teachers and scheduled sessions so they were as accessible as possible, but there was still low participation. Bayview science coordinator Cooper estimated that perhaps several hundred teachers had not undergone formal kit training, corroborated by the fact that more than half of the Bayview teachers who responded to the RSR survey reported as much. And in Benton, even though all teachers interviewed subscribed to philosophies that emphasized the benefits of inquiry and of having students learn through their own experiences, half of the teachers responding to the survey reported that they had participated in only 0–5 hours of professional development in science over the prior two years, suggesting that, in general, participation in the range of professional development offerings was low.

Glenwood offers a dramatic example of the ill effects of teacher turnover while also offering insight into an alternate way to think about critical mass. Due to high rates of retirement and mandatory class size reduction, over 30 percent of Glenwood teachers are within their first three years of teaching. As a result, even in light of a decade of history, program leaders in Glenwood are "at a stage where we just want teachers to open the kit." In the face of this disheartening remark, however, is the fact that Glenwood has survived, perhaps by achieving critical mass of another sort. For 10 years, the district has invested in providing a relatively small group of teachers with professional development—first through the Science Leaders grant and the other early programs and most recently through LSC and USP activities—to develop and deepen their understanding of and commitment to inquiry-based science. Since there is no telling whether teacher turnover in Glenwood will abate, holding to a notion of a critical mass of training for a certain percentage of teachers is ill founded. Rather, when one looks at the capacity that exists in Glenwood, it is clear that the stronghold of sustainability there rests with those teachers and school and district leaders who understand, support, and deeply believe in the value of the program. In them rests the core beliefs and values that will be sustained.

Looking at Glenwood, then, one can explore another perspective on critical mass. In the relative long term, the definition of critical mass expands beyond reaching a critical breadth or number of teachers to focusing on the value of a critical mass of depth or belief among those in a district. If sustainability is "the ability of a district to maintain core beliefs and values and use them to guide adaptations to changes and pressures over time," then critical mass, or "a culture of program self-generation," requires that a sufficient number of teachers and administrators in the district share those core beliefs and values.

Looking retrospectively, those districts that have given attention to ensuring that an even relatively small group of teachers have deep understandings of the goals of the program and the beliefs and values behind them have, albeit perhaps unintentionally, made investments in the abilities of their programs to endure into the future. For example, in Benton, a portion of the professional development under the NSF grant focused on facilitating in-depth learning experiences for school-based leaders. Sycamore also has a similar program for peer-coaches, and Bolton has sought to develop small groups of leaders more than once in its long history. While the intent of this professional development has been, to some extent, to develop the capacity to support other district teachers (a goal that was brought to fruition with variable success), it has also resulted in an important "base" of participation and buy-in to the program that contributed to the idea that hands-on science was accepted practice. Thus, as the programs have matured and evolved, this base of individuals seems to have helped the core beliefs and values remain secure across the district even as instructional materials, professional development strategies, and even leaders have changed.

Sycamore is an interesting district to consider because it is just on the verge of change closely related to critical mass. This district is unique in that it has had dramatically low teacher turnover over the years. In Sycamore, a 10-year veteran teacher might still be considered a newcomer among the 20-and 30-year veterans who have, quite literally, grown up together in their school "families." Many teachers can recall not only the beginning of the current program but also the program's early roots in *SCIS* kits in the 1970s. In Sycamore, there is little doubt that the vast majority, if not all, of the teachers are teaching the science program and that all accept it as the standard of practice in the district. Sycamore has indeed achieved critical mass of breadth.

But, Sycamore faces a wave of retirements looming on the horizon. In response, Allison Stowe is working with the other subject matter leaders to coordinate professional development for new teachers and to focus on kit trainings on a scale she has not had to address since the initiation of the program. The challenge is heightened by the fact that principals also are retiring and the historical trend of hiring long-standing teachers into those administrative positions seems to be subsiding. As the program adapts to the shifting teacher and administrator populations, the notion of critical mass takes on its long-term perspective. The critical mass of breadth is no longer what will hold the program to its core beliefs and values. Rather, it is critical mass of depth—the extent to which the remaining teachers and administrators understand and are committed to the program—that will shape its future.

These data do not suggest that breadth of training is irrelevant to sustainability when compared with depth of belief in a program's core values and beliefs. Rather, these two aspects of critical mass are intertwined, with one requiring more emphasis than the other, depending on where the program is in development. Clearly, breadth contributes to the culture of program self-generation in an ongoing fashion, particularly in the relatively short-term time horizon. However, when programs have experienced shocks, depth of understanding has played an important part in their sustainability.

Montview offers another interesting case to consider. There, after many years of consistent growth and support, the program faced a crisis in the dismantling of the central office and leadership for the program. In the few years that followed, the program seemed to dissolve, leaving only sporadic implementation on the part of individual teachers who used the materials of their own volition. Then, when the district began to revisit and initiate support for a centralized program, the selection criteria for the new program clearly reflected the core beliefs and values that existed in the past.

Another piece of this critical mass picture comes from a study an external researcher did of Montview in 1990, just prior to the collapse of the program. It reports that even though there was evidence of a deep philosophical commitment to the program, evidence for consistent delivery of the program as intended was weak. One can conclude that while actual critical mass of breadth was questionable, it was the critical mass of depth—the commitment to the core beliefs and values—that, though "underground" for a time, emerged as a solid connection to the earlier years of the program.

It is worth noting the close relationship between critical mass and the factors of philosophy (illustrated in the Montview example above), quality, and perception. In the discussions that follow, these factors are intertwined and support one another in overlapping and, sometimes, hard-to-distinguish ways. For example, a look at critical mass through the lens of perception begs the question: *Is it the actual number of teachers teaching the program that contributes to sustainability or, rather, is it the number of teachers who are perceived to be teaching the program*? If looking at critical mass through the lens of quality on the other hand, one might ask: *Is it sufficient for sustainability to have many teachers teaching the program, even if the quality is variable or even unknown*? The answers to these questions (not simple either/or propositions) are discussed more in the sections that follow but are noted here to acknowledge and illustrate the interactivity of the factors contributing to sustainability, particularly those that pertain to the whole science program.

In summary, the data of the study suggest that the conventional notion of critical mass, i.e., "If one reaches a critical mass of teachers, the program will be sustained," is largely a myth. Even in those places where a vast majority of teachers are reached, other forces such as changes in district leadership, increased accountability, and shifts in financial resources can pose potentially insurmountable challenges. An alternative proposition is to consider whether an emphasis on breadth of critical mass (reaching numbers of teachers) or depth of critical mass (belief in and commitment to the core beliefs and values of the program) will help to withstand these and other shocks to the program at various points in its evolution.

PERCEPTION

- *The perception of a science program can differ greatly from the actual status of that program in a district. "Misperceptions" can both contribute to or inhibit the sustainability of a program.*
- *In the absence of firsthand knowledge of the status of the program, program leaders and other decision makers take action based on their perceptions.*
- *There is a disconnect in perceptions of the status and importance of the program held by stakeholders at different levels. This confounds efforts to accurately diagnose and address needs.*

Perceptions—whether held by program leaders, program participants, or outsiders to the district—significantly supports and inhibits the sustained programs. In some cases, perceptions of the programs differ greatly from the apparent actual status of the program. This is significant because, in the absence of enforced accountability measures, perception becomes 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 strong and exemplary. While this impression is positive, it also opens the door for potential neglect in allocations of future district dollars and attention.

Looking at the older programs, for example, positive district-level perceptions were key to program establishment. In the early years in Bolton, for example, Pearl North garnered support from the school board by demonstrating the use of kits for them and courting the board members' interest over time. Although most board members never saw the program in action, their enthusiasm and sense of its value to students stemmed from the impression that North had created. In Hudson, Lawson effected a similar outcome with the school board, as did Donahue with his "clay boats" activity used with the board in Montview. These early and continuing efforts to bring school board members "into the fold" developed in them a view of the programs' worth that contributed to their willingness to support them over time.

As these earliest programs became established, they developed national reputations outside of their district, as did their leaders who were considered to be ahead of their time. "Mythologies" developed that reinforced the view at home that the leaders and their programs were rare and valuable, which, in turn, served as protection for the program. Lawson was a charismatic person with seemingly boundless energy, and stories of her Herculean efforts to build the program abound. Moreover, two legends of the impact of Hudson's science program still circulate. These include one story of the year that an overwhelming number of sixth grade students from a "kit program" elementary school chose to enroll in science at their junior high compared with almost none of their peers from a "textbook science" school. The other story focuses on a letter that arrived in the superintendent's office from Stanford University saying that there were more acceptances for the new freshman class from Hudson than from any other school district in the country. Both Hudson and Montview attracted national attention in media outlets, and all of the early leaders received national recognition as pioneers in their field.

Though the national and local recognition focused value on the science programs, taken together, they also tended to foster a kind of complacency in these districts, where the perception reigned that the science programs were exemplary and secure, even when there was evidence to the contrary. In Bolton, following the close of the LSC grant, administrators were asked, "What happens when science isn't taught?" Their response was incredulity; their understanding is that the purpose of the LSC was to promote the program through professional development. Since the money has been spent, they believed that the program was being taught, even though the evidence suggested that implementation was still uneven. In Bayview, one superintendent commented that, "Because science kits are so much a part of our culture, it has not been a discussion topic since I have been here." Even in younger programs, where the track record is not as long, the notion that the program is secure and in good shape can lead administrators to be complacent. The deputy superintendent in Garden City observed that "When it's not good, I hear from the parents. So no one is currently complaining about the program, and science is not controversial, so it is not on the radar screen." In fact, the program may suffer from what researchers referred to as "lack of squeaky wheel syndrome," meaning that because no one is raising concerns about or objections to the program, it may not be getting the continued attention, resources, and support it needs.

Another contributor to this complacency is the endurance of the program itself. As programs gained maturity while avoiding significant conflicts, they became accepted as district icons. As the director of evaluation in Garden City observed, "…There is a point at which inertia takes over: what do you mean you are going to do away with this? We have always done this." The first science consultant there remarked, "It has been sustained because other people see it as a success. And because they have had success with it, they want to keep it going." Thus, staying power—the mere fact that the program has been in place for a marked amount of time—can contribute to a program's sustainability, not as a result of demonstrated quality but, rather, as a result of how others interpret the program's endurance.

This phenomenon, which one might characterize as sustainability through passivity, has helped these programs grow and develop over time, but it has also contributed to a lack of critical investigation of the program, adding to the already challenging responsibility of program leaders to understand and provide appropriate supports. As the discussions of accountability and quality illustrate, the capacity of program leaders to assess the actual status of their programs is minimal at best. Voluntary professional development with no means to evaluate its impact on classroom practice and minimal capacity to assess the depth and breadth of instruction across the district all leave leaders with only a limited understanding of their programs. Thus, in the absence of empirical data about the status of a program, and with the notion at the administrative level that it is a success and the accepted way of doing things, these hands-on programs have largely been sustained on the strength of leaders' and administrators' perceptions, regardless of whether or not they are accurate. This could be seen in Montview, where a case study of the program was written by a district outsider between 1988–90. In it, the observation was made that teachers were using the kits to a much lesser degree than expected. Regardless, the program was fully institutionalized at the district level, and administrative decisions reflected their belief in the program's sound condition.

When science flies below the radar screen, the lack of attention allows programs to proceed without pressure. When controversy attracts the attention of the central office administration, that freedom is in jeopardy. In Bayview, program leaders presented the new science content standards that accompany the state academic requirements to the school board and the superintendent, leading them to begin paying more attention to science than before. They asked questions that had not been asked in years, such as, "Does everybody teach the units? What are the statistics?" This increased attention is an important shift from the complacency described earlier. Instead of being relegated to the "back burner" because a program is assumed to be going strong, being in the spotlight holds both positive and negative implications for the science program. Increased priority may be good, but the increased inspection may be a challenge to contend with.

The positive perceptions of central office administrators are critical for support of the program at the district level, but implementation of the program depends on the views of teachers and principals. As the educational leaders for their school, principals have tremendous influence over what is taught in their buildings, and teachers are very aware of what their principals expect and value. Teachers are also influenced by their own perception of the program's value, while principals are persuaded by the expectations and values of the district's administration. Thus, perceptions of the science program at different levels of the district have different implications for a program's sustainability.

Principals and teachers in this study look to the central office for direction regarding how to implement the educational program within their schools. One principal commented, "Foremost is the commitment of the district to the program…from the top on down…it has to come from the top…" In Sycamore, the cohesiveness of the district administration and the support they feel for the science program has been easily conveyed to the schools beyond the administration building. In Garden City, teachers also commented on the importance of support and commitment from "downtown" and feel that the central office is behind the science program. Similarly, the central office administrators perceive the teachers as being fully on board with their implementation of the program. In Montview, as the program was first being established, the program coordinator believed that the support of the board was crucial. By establishing guidelines for science instruction, they conveyed the expectation that it would be taught.

For their part, it is not uncommon for teachers to express a sense of being "valued" by program leaders, primarily as a result of the ways leaders cultivated teacher leaders and their professional development efforts. Leaders in Bayview, for example, engaged teacher leaders by involving them in the program, providing them with additional professional development; and including them in the development of curriculum, unit assessments, and the like. Teachers responded with a growing commitment to the program that, as one teacher said, "listened to the teachers." In Benton, teachers are similarly loyal because they feel the program leader respects their opinions and responds to their concerns. During the 1980s in Montview, classroom teachers received support for teaching science through a program called "Comfort and Caring," in which lead teachers would observe a science lesson and give teachers feedback and support. Although these visits were relatively infrequent, teachers indicated that this was a very strong and valuable component of the science program; it generated a great deal of respect and loyalty because teachers felt they were understood and respected. Although it is difficult to say with certainty, it is reasonable to conclude that this kind of loyalty and appreciation will increase the likelihood that teachers will teach the science program.

In contrast, teachers' perceptions of the value their principals place on science teaching are variable but generally low. In the survey of principals and teachers conducted for this research, teachers in every district responded that they think their principals place a lower value on science than the principals actually feel themselves. Likewise, principals also underestimate the importance that teachers place on science. Further, only about 35 percent of the teacher respondents feel their schools' administration strongly supports science, while about 77 percent of the principal respondents feel they are strong supporters of science. When considering the frequency with which principals observe science instruction, it is easy to see why teachers would feel science is below their principals' radar screen. Only about 16 percent of the teacher respondents have been observed teaching science in the current year. This perceived lack of support for science, whether accurate or not, suggests that teachers who feel a commitment to teaching science may also feel challenged to do so by a principal who is implicitly or explicitly unsupportive.

The mixed messages seen within schools return this discussion of perception and its impact on sustainability to the central point. Given the lack of authentic data on the status of a program, perceptions of it are often all that decision makers have to guide their actions. The fact that there are disconnects and misperceptions at every turn make the challenging job of growing a districtwide science program even more difficult. It also suggests that perception has been sufficient to sustain these nine programs up until now. In an environment of increased scrutiny, however, it is impossible to say whether perception alone will continue to be adequate.

PHILOSOPHY

- *In sustained programs, there is a widespread, shared philosophy that science should be taught using a hands-on approach.*
- *Science programs become vulnerable in the presence of inconsistent philosophies about the importance of teaching science.*
- *The growth of the hands-on philosophy is supported when there are pre-existing or newly emerging complementary approaches elsewhere in the district.*

In this study, philosophy refers to a set of beliefs about the role of and appropriate pedagogy for science in elementary education. Teachers, principals, administrators, and program leaders all articulated "philosophies," although it is unclear whether the practitioners themselves would feel comfortable with that formidable label to describe their views. They might prefer a less formal one and, indeed, the euphemism for culture suggested earlier —"It's the way we do things around here"—could be adapted in this case to philosophy—"It's what we believe around here."

Across districts, expressions of belief regarding science instruction have grown out of practitioners' understandings of how children learn in general, and how they learn science in particular; their views about the most successful methodologies for engaging children in the learning process; and their beliefs about environments and approaches that foster children's enjoyment of learning. These beliefs were expressed through a range of channels. For example, the assistant superintendent from Bayview stated, "I believe, as do a lot of people now, that the way to teach science is to have kids engage in it… By getting involved they learn the process and the content of science. That's what the kits are all about, in 25 words or less." An example of a written statement comes from Hudson's Science Content and Performance Standards: "The foundation of solid scientific understanding should be based on a rigorous exchange of student- and teacher-generated questions and responses, supported by experiential classroom activities under the guidance of a knowledgeable instructor." Finally, the comments of two teachers from Benton typified philosophy statements made by teachers. First, "My philosophy is that kids learn way more when they are generally curious, when something real is happening, and when they can ask their own questions," and "Science gives my kids the chance to experience the joy of learning."

This study demonstrates that philosophy, as it was expressed by teachers, principals, and administrators in the sustained programs, falls into two categories: (1) beliefs about the importance of teaching science, and (2) beliefs about how science should be taught. These two philosophical strands evolve, sometimes together, sometimes independently. In sustained programs, the second strand, relating to how science should be taught, is consistently strong—educators in these districts articulate beliefs that the hands-on approach to science instruction is the best way to teach science. However, the first strand, representing belief in the importance of teaching science at the elementary level, fluctuates depending on the changing district conditions. Thus, even though the programs demonstrate widespread common beliefs about science instruction, they remain vulnerable when lacking support for making science a core part of the elementary student instructional experience.

Looking more closely at the first strand—the importance of teaching science—none of the teachers, principals, or administrators made statements suggesting that they do not believe that teaching science is essential. In fact, many administrators across the sites expressed support for the importance of science teaching. At the same time, however, there is a great deal of willingness to forfeit science instruction in favor of reading and math, particularly when high stakes tests are in play. Most principals, for example, declared their support for science, but their behavior belies their declaration, in that they generally have only cursory knowledge of the science program, make infrequent observations of their staff's science instruction, and make little or no attempt to encourage reluctant teachers to use the science kits.

Central office administrators often exhibit the same conflicts. They may have affirmed that science is important but exercise little oversight of building principals to ensure there is some accountability for program delivery. The lack of a clearly articulated view that instruction in science is important, hand-in-hand with a lacking or absent accountability system, increases vulnerability and even threatens the sustained program. In Hudson, for example, an evaluation report of the LSC suggests that many teachers and principals share the belief that mathematics is a primary curriculum focus at the expense of science. As a result, many teachers feel unable to include science regularly, if at all, because they feel the need to emphasize mathematics and reading. In some cases, teachers return kits to the materials center partially or completely unused.

Another example comes from Glenwood, where the belief in teaching science using hands-on approaches has been strong since the beginning and is espoused through writing and messages conveyed in professional development experiences. This point is clear, but it is dramatically tempered by the fact that it is not accompanied by strong verbal and nonverbal messages about the importance of teaching science. Simply put, principals still are held accountable for reading scores. As one central office administrator stated, "Principals are not 'against' science…they are just 'overwhelmed'…the district priority is literacy." Some shifting of priority in Glenwood has come with the arrival of the USP grant that requires accountability systems for science with possible sanctions for low performance. But this raises a complex but important concern: Is there actually a philosophical commitment to the importance of teaching science? Or, is that commitment simply defined by the presence and/or absence of accountability systems for science instruction? One would hope that the commitment would exist independent of the accountability system, but this, of course, is an ideal not observed in these programs.

Montview and Bolton offer examples of how the pressures that reveal weaknesses in the commitment to science instruction can happen numerous times in the lifetime of a sustained program. For example, in the mid-1970s, Montview's science coordinator Thomas Donahue wrote, "The strong emphasis on reading, mathematics, and language arts in the last several years at the elementary level apparently has reduced the time and effort given to science in many schools." Thirty years later, vulnerability of the philosophical commitment to the importance of science instruction is evident once again. One of Montview's professional development staff remarked that principals are held accountable for student performance only in literacy and mathematics, thus, it is "easy" for teachers to feel science is less important, and even neglect science entirely. Similarly, Bolton experienced a program decline in the 1980s that program leader Pearl North attributed, at least in part, to the emphasis on reading and mathematics standardized tests. The director of curriculum and evaluation explained that because these tests do not include science, they divert attention away from the program and lead many teachers to lose their commitment to science teaching.

Educators spoke frankly and with some regret about the need to devote most of the school day to instruction in math and reading, suggesting that indeed, they do feel some obligation to science instruction. That obligation, however, is not always bolstered by accountability systems or dictates from the central office leadership. At these times, when districts show some weakness in their commitments to the importance of science teaching, the sustained programs are most vulnerable and, in some cases, recede into holding patterns or even move underground until renewed interest in science creates a more hospitable environment for new growth and evolution.

The second strand of philosophy, quite distinct from the first, focuses on beliefs about how science should be taught when it was taught. Here, variation within and across districts is greatly reduced and, in general, most districts appear to have cultivated a strong common philosophy about how science should be taught. It is worth noting that even in the programs that were relatively young, such as Sycamore and Benton, the groundwork for the commitment to a hands-on approach to science instruction has been laid by early use of kits in a more sporadic or informal way. The focus on using materials and the importance of students actively doing science rather than passively observing it is quite consistent.

In Bolton, for example, a regular adoption cycle occurred at about the 10th year of the program, prompting an examination of the kits and the curriculum. This examination revealed the need to update the kits, and set in motion a two-year process that resulted in redesigned kits with, among other things, more carefully conceived activities, targeted teacher's guides, unit objectives, and student assessments. Such a wide-scale revision process would not have been possible without a shared belief in the philosophy of the program across all levels of the district, from the classroom teacher to the program leader to the central administration.

In Garden City, on the other hand, when the state standards in science brought on an examination of the science program, the program leader implemented adaptations that included changing the grades in which specific topics were taught, buying new kits, and including the use of textbooks to address topics not covered by kits. At the same time that the inclusion of books suggests a weakening of commitment, teachers' comments about the textbooks suggested that the possibility they would supplant the kits was not a concern. They simply viewed it as a means for ensuring that, in this district driven by the state standards, they would be able to cover all of the material. Here, revising the kits as they did in Bolton, was not a viable option and the pressure to respond to the standards was 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 is evidence of and reinforcement for the importance of the use of materials to teach science. Here again, it was the shared commitment across the district that made this adaptation succeed.

A dramatic example of commitment to a philosophy can also be found in Lakeville, where after 15 years in place, the district's elementary science program was suddenly in serious jeopardy. A shift in political agendas on the school board created a public debate about its value, and a public show of strong support for the program 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, each expressing their approval of what the science program had achieved for the district's students. The program survived intact, but only because there was widespread understanding of and appreciation for the science program's philosophy among the public. 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—not the result of hastily coached volunteers but of those with personal experience who had seen what the program had achieved since it began.

Finally, Montview presents the most extreme example of philosophical endurance. After more than 30 years of strong development and evolution, central support for 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. By most definitions, the program disappeared and, in fact, it did become invisible at the district level. However, the data collected in this study suggests that rather than disappear, the program went "underground," and those teachers committed to science continued to use the kits to provide science instruction to their students. Teachers and administrators who were there throughout this period refer to it as the program's "dark time," but although the visible artifacts of the program have disappeared, their loss should not lead one to presume that the belief in hands-on science died as well.

Within several years, the need for centralized functions re-emerged and the centralized science program began a revival that focused, once again, on using a kit-based approach. In fact, as the new curriculum was being planned, the commitment to the same core philosophy was never in doubt, and traditional textbook programs were never 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, there was enough activity and commitment to retain the institutional memory of the importance of the hands-on approach. After several years of neglect, the program's new leaders did not need to start from scratch–the years of an earlier program established an elevated understanding of the philosophy supporting kit based programs and formed a foundation upon which the new program was built. Here, even when not part of the district structure, the understanding and acceptance of the hands-on approach resided in the hearts and minds of teachers and principals who embraced its revitalization.

Some sites produced evidence that compatible philosophies in other subject areas may bolster the development of the commitment to hands-on science. Sycamore provides the most dramatic example in that, prior to program establishment, a number of teachers and principals participated in the Follow-Through program, which emphasized the importance of student-centered learning. Another example comes from Bolton, where the interest in the *Kagan Cooperative Learning Program* meshed well with the emphasis on hands-on science instruction, thus, reinforcing them both. In Hudson, the way was paved for the use of science materials because a similar approach was used in math and art. A materials distribution center for these subjects was already in operation, so it was a short conceptual leap to establish a similar system for science.

Still, it is worth noting that the presence of a shared general philosophy does not suggest that classroom instruction is consistently aligned with that philosophy. As discussed in the section on quality (p. 89), classroom practice observed in this study was highly variable. Some teachers engaged in instruction that reflected the generally accepted philosophy about science teaching in the most minimal sense; others used the kits provided but in a very mechanical way, leaving little room for the flexibility that often accompanied skilled teaching. And, still others' instruction reflected a more sophisticated understanding of inquiry that most closely matched the formally stated goals of the program. Still, regardless of the variation in actual practice, the bedrock acceptance of doing science with science materials was sound.

Although the two strands of philosophy are related, they are not mutually dependent. It was not uncommon to find districts where the commitment to teaching science had varied greatly over time, while the belief in teaching science with kits remained strong. While teachers and administrators professed to believe in the importance of teaching science, they admitted that when faced with other priorities, science could be neglected. At the same time, there was striking consistency in instruction; when science was taught, there was widespread agreement it should be taught using "hands-on" activities or "inquiry" approaches. In the face of a focus on other subjects, the science programs sometimes adapt, adjust, hold their ground, or even retreat somewhat, while still holding fast to this second strand, the importance of teaching science using a hands-on approach. The belief in the importance of teaching science must be extremely strong to withstand the pressures that come from accountability for other issues. Generally, that belief has been strong enough to sustain programs through challenges, but not necessarily strong enough to give the sustained programs a sense of security. Only when both strands are strong does program vulnerability fade.

Quality

- *There are no effective mechanisms in place for assessing the quality of science instruction and/or the impact of professional development.*
- *In the absence of accountability measures, actual student learning of science concepts and processes becomes irrelevant to a program's sustainability.*
- *In the presence of accountability measures, program quality is defined by evidence of student performance on those accountability measures. Thus, the degree of alignment between the program and the district's accountability system becomes the primary indicator of program quality.*

Although it was not the intention of this study to assess the quality of any program, it is reasonable to ask whether quality has a role in a program's sustainability. Thus it is important to have a working definition of the term, "quality." This study defines the quality of a program as the extent to which its instruction and curriculum 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. This view of quality stresses two inputs—instruction and curriculum—and it points to two outcomes—student attitudes toward science and student understanding of science process and concepts. With that definition in mind, it is possible to have a high-quality program using any number of instructional approaches and curricula.

If the quality of curricula and instruction is to have an impact on a program's sustainability, there must be mechanisms in place that allow program leaders and others to gain and maintain an understanding of their status. What is most striking, in this regard, is that none of the districts in this study have any such systems in place. It is impossible for any of the program leaders to have a sound understanding of the quality of instruction or the impact of professional development on classroom instruction.

This situation is the result of several of the factors and conditions already discussed. Most pertinent, for example, is the fact that program leaders have virtually no decision-making authority or power beyond the components of their programs. They are not the teachers' supervisors and, thus, have no authority to go into classrooms and observe instruction unless they have been invited to do so. This necessarily truncates their view of their programs' status. Moreover, though some districts have resource teachers who spend time in classrooms, none have enough on staff to maintain awareness of classroom practice districtwide. This is particularly true for Bolton, for example, where two program leaders and no resource teachers serve about 60 elementary schools. Additionally, even if there were enough resource teachers, they don't want to—nor do their fellow teachers want them to—play the role of "hall monitor" in the classroom.

Hudson has arguably the greatest potential for maintaining an awareness of classroom instruction as a result of the many forms of support available for teachers. As described in the site report, these include resource, liaison, connection, and mentor teachers, as well as instructional specialists. It remains to be seen what they will be able to accomplish across classrooms when the main concern has been addressing the immediate needs of novice teachers. Moreover, the liaison, connection, and mentor teachers are all vulnerable to the vagaries of grant funding, so their contributions to long-term sustainability are hard to predict. Finally, the supports in Hudson described above are a recent addition, with the exception of the resource teachers who, for about 20 years, have provided the training and assistance for classroom teachers. This has been the case for most other programs and continues to explain their limited capacity.

Limited capacity notwithstanding, leaders expressed frustration with their inability to maintain an ongoing understanding of the status of their program because they know the importance of this knowledge. In 1995, for example, as Lakeville's NSF grant was concluding, program leader Wolters was asking questions of the classroom teachers, such as "You have been teaching this [program] for five years now. Do you know what the main points of the lesson are? How do you know if the kids are learning any thing? …What is it that you are not doing that is not helping the kids get to the next level?" Without district support for the full panel of resource teachers, this kind of follow-up support is possible for only a small number of teachers. Moreover, resource teachers' contact with teachers and their principals was reduced, and they could no longer keep their fingers on the pulse of what was happening in science lessons.

Over the course of their programs' history, several leaders have made attempts to understand the status of their programs, and their findings corroborated the findings of this study: implementation of kits within each district is uneven, and, when teachers do use kits, their practice is highly variable. In Hudson, for example, teachers who responded to the RSR survey reported teaching science for an average of 116 minutes a week, which is very close to the district goal of 120 minutes. However, this is inconsistent with focus-group discussions with experienced teachers conducted as part of this study, as well as a survey conducted in 1999 by the leader of the program at that time. That survey, distributed to 430 elementary teachers in 15 of the 50 schools, had a return rate of 95 percent. In it, 77 percent of the respondents reported teaching science for 90 minutes or less per week, while RSR focus group participants reported teaching science for "about 60 minutes" per week.

Hudson's story is common among the nine districts, both at present and in the past. In Glenwood, for example, just over half of the respondents to the RSR survey reported that they are expected to use four kits per year (the district's requirement), but less than 10 percent reported that they actually use that number. Further, only slightly more than 10 percent reported that they had been trained to use the four kits they are expected to teach.

In Montview, the program had been in place for about 10 years when an external researcher conducted an evaluation of the implementation of the program in 1979. The report concludes that "…science is not a high priority instruction area and cannot compete with the district emphasis on teaching reading, language arts, and math. Science instruction is often viewed as … a 'frill' but never a basic skill on which life-long learning will be based." The report goes on to say that, in a random sample of 11 out of 44 schools, 44 percent were not teaching the science program for the recommended amount of time. These findings agree with the common view expressed by teachers who were interviewed for this project. Across districts, they reported that the emphasis on student achievement in reading and math has limited the amount of time they can spend teaching science and, as a result, they are teaching less science.

With uneven implementation as a backdrop, the instruction that is provided also varies widely within districts. Again, evaluating the quality of instruction was not the aim of this study; however, the classroom observations served an equally important purpose. RSR researchers asked to see instruction that matched program leaders' views of the kind of teaching that they wanted to promote and thought was realistic to attain across the district. In other words, RSR researchers did not ask to see the "best" or the "worst" teachers, but a sample of instruction that represented what program leaders thought they could achieve. RSR researchers observed about 20 classrooms in each district, with teachers ranging from novice to veteran, and schools ranging in SES and demography; observations commonly included a whole lesson and lasted about an hour.

Overall, the observed instruction could generally be described as "mechanical kit use" and, within that broad category, there was considerable variation in several dimensions. First, teachers varied widely in the way they used the materials. Some reported using the kits from start to finish, but this was the most uncommon strategy, usually because of time constraints. More frequently, teachers selected specific activities from the units, circumventing the sequence of learning activities that had been planned for the unit, thus reducing the units' intended goals of developing understanding of science concepts. Of those teachers who responded to the RSR survey, 48–78 percent (depending on the district) reported that they "pick and choose" parts of the units to teach.

Second, there was variation in the extent to which lessons were introduced and students' prior knowledge was elicited. Some engaged their students in a full discussion, recalling students' previous work, concepts, and vocabulary, while others were very brief, spending most of the opening minutes giving instructions so students would conduct the day's activity according to the teacher's expectations.

Third, on rare occasions, students generated their own questions regarding what aspects of the phenomena they were going to explore. More commonly, they either proceeded according to their teachers' prescribed set of instructions or chose an experimental approach from prepared options.

Finally, teachers concluded their lessons very differently. It was most common for teachers to wrap up their lessons quickly and instruct students to clean up so they could move on to the next subject. Only occasionally did a teacher help students connect the days' lessons to the "big ideas" of the unit. These observations were quite contradictory to the data collected on the teacher surveys, which demonstrated that, in every site, between 78–90 percent (with the exception of one that had 68 percent) of all responding teachers reported that they "lead a discussion in which students talk about the meaning of what they have done."

Within this general framework, there was also variation in teachers' abilities to focus a lesson, explain the planned activities, ask insightful questions that elicited thoughtful responses, manage the materials, and make sense of the activities with their students. Some teachers felt very comfortable with the science concepts and materials, while most were clearly out of their depth. This variety in instruction extended to the rooms' decor. Some rooms had an abundance of student work in science on display along with artifacts, and science equipment, while other rooms showed little evidence that science was taught. Notably, there was no correlation between the impressions the artifacts in the classroom conveyed and actual classroom practice.

In the same manner that leaders are unable to maintain an awareness of the breadth and quality of instruction, they are also ill equipped to assess the impact of the professional development they provide. Combining this challenge with the high rate of teacher turnover in some districts and the voluntary nature of professional development in all districts suggests a discouraging outcome. As a whole, these nine programs probably do not have the spread across their districts that a "districtwide" program would suggest, and moreover, when a program is delivered, the quality of instruction is inconsistent and, most likely, not as close to program leaders' expectations as they would believe.

As disappointing to program leaders as these findings may be, considering the effort and resources they devote to professional development, it implies something quite unexpected for sustainability. The fact that this condition of incomplete knowledge has been the state of affairs for all of the sustained programs in this study since their inception suggests that the quality of instruction and professional development appears to be irrelevant to a program's sustainability. The importance of professional development to sustainability is discussed in the previous section on professional development (p. 65), and its contribution is valuable, but surprisingly not necessarily because of its impact on classroom instruction.

Each program in this study came into being at a time when little attention was paid to student outcomes in science. With the exception of Garden City, standardized tests in science were not in use five years ago (Montview administered science tests early in its program), and as a result, student learning of and attitudes toward science were not taken into consideration when central office administrators estimated the quality of their hands-on programs. Because student achievement was not considered, neither was the quality of the curriculum, and since the breadth and depth of instruction is largely unknown, these factors are also immaterial to administrators. Instead of including instruction, curriculum, and student outcomes in their "quality" formula, other features come into play.

Administrators in Garden City, for example, told of their satisfaction with the program because there are no complaints from parents or from teachers. Others have seen the kits in use and are impressed by them, because the materials centers function well and teachers express their appreciation for the service the centers provide. Administrators assume that if all is quiet, then all is well. Thus it is reasonable to conclude that before high-stakes accountability measures came on the scene, instruction, curriculum, and student outcomes neither contributed to nor inhibited the sustainability of these hands-on science programs.

Since the recent trend toward high-stakes testing has taken hold, however, tacitly accepted measures of quality have changed. Gone are the days seen in Hudson, when poor student performance on a science test was considered a result of a poorly designed test as opposed to a poor science program. Now, student performance on these tests is the accepted evidence that the program is accomplishing its goals. In this environment, where state tests are intended to reflect state goals, program leaders now look to the national, state, and/or local science standards for guidance. As discussed in the section on adaptation (p. 70), program leaders in almost every district in this study have undergone a review of their curriculum and adjusted it to improve its alignment with the standards in preparation for standardized science tests. They have come to use the degree of alignment with their science standards as their measure of the quality of their curriculum.

Debates continue in Benton, Sycamore, Bayview, and Garden City about the quality of the state science tests that have been or are being developed. Estimations of the degree to which they accurately reflect the science standards differ, but the outcomes of these debates are of minor importance while the public emphasis continues to be on student performance. In Bayview, for example, the only way to prepare continues to be through alignment. In Garden City, Sycamore, and Lakeville, the inclusion of science textbooks in the curriculum is seen as a way to ensure that all topics are covered. The fact, too, that the public eye is turned most critically to student outcomes in language arts and math does not relieve the pressure. Although state tests are in place in Benton and Sycamore, there are no consequences associated with low science scores, likewise in Glenwood and Bolton. Regardless of the low priority administrations and the public have placed on science, leaders feel the need to have some evidence on hand of student achievement. This need fostered several different strategies among the nine districts.

In many districts, such as Bolton, Bayview, Benton, and Lakeville, considerable work has been done to develop student assessments that teachers may use with each unit. However, as program leaders in Bolton attested, classroom teachers rarely use their assessments for this purpose. Furthermore, even if they were used as intended, the district does not have the capacity to analyze the data to understand student mastery districtwide. In response to the need to have some data on student performance in science, the district has added the science portion of the California Achievement Test (CAT) to its testing portfolio, even though it is not aligned with the curriculum. The effect that student scores will have on the science program is unknown. Hudson employs the SAT9 test in science, as a tool to monitor the science program, also not aligned with the curriculum. Students appear to do well on the test, so little comment is made. In Lakeville, the effort to develop unit assessments was also unsuccessful, and the lack of student outcome data left the program very vulnerable to public criticism of its quality. In Glenwood, the district developed its own science test and program leaders are pursuing mechanisms to increase the importance of this test as a way of increasing attention to science overall.

These findings regarding the definition of quality and the role that it has played in sustainability are poignant. Until the recent past, central office administrators and the general public placed relatively little emphasis on the elements of quality that today are considered all-important. Regardless of the investments that leaders made in the components of their programs, student learning, instruction, and the curriculum were relatively insignificant as long as the program was seen to function smoothly with no complaints.

Since the importance of student achievement on standardized tests has taken hold, the definition of quality has come to mean student scores on science tests. The pressure exerted by these tests has driven leaders to reshape their programs; however, the factor that most directly influences student achievement—instruction—is still outside the view of quality. Thus investments in professional development remain minor, particularly compared with professional development in reading and mathematics instruction. Moreover, as test scores take hold as the measure of success, the possibility of bypassing hands-on curriculum in favor of textbooks becomes more attractive to teachers and principals. The implications for future sustainability are worrisome.

CONCLUSION

> What the textbook can summarize in a page of results—life is cellular, cells have water and carbon, cells divide to multiply—our methods with the child's own work, with his own hands, with his own microscope and labored arithmetic may take six weeks of classroom effort… We are not disturbed by slowness, for what goes slow can run deep. And school hours are not all of life. To stroll into reality, the detail of it and the context, to unravel and uncover, is a better thing than to sprint past, reading the billboards of science.
>
> —*Philip Morrison, Physics Professor, Massachusetts Institute of Technology, Co-Leader of ESS*

This quote embodies the essence of the hands-on science instruction that inspired the programs studied in this research, and conveys deeply held beliefs about how students should learn about and interact with the discipline of science. Education leaders across the country have subscribed to this philosophy and dedicated themselves to building and growing elementary science programs powered by its implications. And yet, in spite of the fervor with which they are committed to these beliefs, their programs remain vulnerable. In an effort to understand how to reduce this vulnerability and thus increase the capacity for hands-on science programs to endure, this study focused on nine programs that have done so, for from 10 to more than 30 years. The research team collected data in these districts for three years to answer the question: *What contributes to or inhibits the sustainability of a districtwide, hands-on inquiry science program?*

Throughout this research, program leaders expressed the hope that a consistent pattern would emerge from the data collected across these nine programs and offer a formula for sustainability that would guide their efforts. They dearly wanted more knowledge about how to maintain their programs, strategically concentrate their efforts, and build capacity for continuous growth and improvement. However, as evident in the preceding discussion of the findings, no such formula emerged. Rather, this study identified a set of factors that affect the sustainability of hands-on science programs in fluid and interrelated ways. The roles these factors play in reform efforts are greatly varied and change over time and from place to place as they reflect the complex school district environments around them. Within this complexity, while there is no formula for sustainability, the factors presented here illustrate trends that offer new insights into sustainability for program leaders, district administrators, and funders as they invest in new and ongoing reform efforts.

THE FACTORS

The factors that support and/or inhibit sustainability of districtwide hands-on science programs fall into three categories: those that pertain to conditions surrounding the district and its program, those that pertain to individual components of the science program, and those that pertain to the program as a whole. See Table 2 below.

Table 2

Factors that Pertain to Surrounding Conditions	Factors that Pertain to Science Program Components	Factors that Pertain to the Science Program as a Whole
Culture	Accountability	Adaptation
Decision Making and Power	Implementation	Critical Mass
Science for All	Instructional Materials	Perception
	Leadership	Philosophy
	Money	Quality
	Partnerships	
	Professional Development	

It is critical to recognize that in interpreting these factors' impacts, one must disregard the linear nature in which they are presented above, and understand that the borders that separate the categories and the factors themselves are porous and frequently shift. The site reports offer illustrations of these interrelationships through examples of cultural influences on leadership, ways philosophy emerges in professional development, and the tension between accountability and the notion of science for all. The possible exchanges between, and connections among the factors are innumerable—in fact, this flexibility lies at the very core of how and why these findings offer insights and relevant implications for program leaders, district administrators and funders, all of whom operate in unique, idiosyncratic environments.

To a great extent, the factors that pertain to components of the science program—accountability, implementation, instructional materials, leadership, money, and professional development—confirm much of what already is known. Those experienced in science education reform would expect that these factors would play an important role in sustained programs, and in fact, they all were included in the initial data collection framework developed as part of the study. Indeed, they are unquestionably critical to sustainability but in contrast to their expected significance, they often demonstrate their import in unexpected ways.

The factors that pertain to conditions surrounding the district—culture, decision making and power, and science for all—though not typically addressed in discussions of sustaining reform, also were not completely unexpected. As surrounding conditions, they create the milieu, or set the stage, for the drama of science program establishment, maturation, and growth. They rarely were explicitly articulated by individuals in the sites, nor do they have a highly visible presence in district or program reports or documents. And yet, though they often are subtly embedded in the data, their impact and influence is without question.

The factors that pertain to the science program as a whole—adaptation, critical mass, perception, philosophy, and quality—are the least predictable and, in some cases, surprising in the ways they do and do not contribute to the programs' endurance. Program leaders did not necessarily know that these factors have played a part in their programs' long lives, nor did they know that they as leaders have had a hand in developing these factors over time. Nevertheless, as evidenced in the histories of the older programs, these factors are key influences on sustainability. The decisions and investments that program leaders make directly affect the extent to which these critical factors positively and/or negatively affect the sustainability of the program and the strength of that effect at any given time.

THE FACTORS AND THE PHASES

Having laid the groundwork with an overview of the factors, it is worthwhile to return to the phases of a program's development described earlier in this report: establishment, maturation, and evolution (p. 21). Typically, program leaders in the establishment and into the maturation phases are focused on the second group of factors—those that pertain to components of the science program. They are intent on launching their materials centers, organizing professional development strategies, identifying leaders, and securing funding—all vital to a healthy program's beginning. One might consider these factors essential to what RSR researchers refer to as "short-term" sustainability—meaning, sustaining the program through the first several years—throughout establishment and into maturation.

As a program matures, continued attention to these factors remains critical but should not monopolize the program leaders' attention. An exclusive focus on the factors related to the program elements overlooks the significance of the third set of factors—those that pertain to the whole program. Much as the factors that pertain to the components of the science program affiliate with short-term sustainability; the factors that pertain to the whole program affiliate with "long-term" sustainability—meaning, sustaining the program for 10, 20, and even more than 30 years. Without attention to the third group of factors, a program has limited support for making a transition through maturation and into evolution—or, in other words, into the domain of long-term sustainability.

Having discussed the phases and their associated factors, it is important to note that the distinctions made between phases in actual programs are not so discrete. The development of the phases overlap and are never left entirely behind. Even when a program is mature, if a new superintendent joins the district or new teachers and principals are hired, the program leader must re-visit establishment to ensure the program is once again introduced, understood, and accepted. Similarly, even as a program evolves, with each major adjustment, program leaders must return to maturation to ensure that the program, as it is newly interpreted, continues to be routine and accepted as standard practice. As a result, leaders of sustained programs must simultaneously engage in activities that support both short- and long-term sustainability concerns.

IMPLICATIONS

These findings offer many implications for program leaders, district administrators, and funders with regard to their investments in their science programs. Some are described below:

- *Leaders and supporters of districtwide programs can gain from giving attention to the wide range of factors that affect sustainability and account for them in all strategic and financial decisions.*

The leaders of the sustained programs in this study emphasized the factors related to program components throughout the lives of their programs—even as they moved out of the establishment phase and into maturation and evolution. Thus, their investments in and accounting for the other factors have been a fortunate by-product. This study serves current and future program leaders by making the factors that pertain to surrounding conditions and the whole program more explicit, allowing leaders to be more purposeful about how, and in what ways they allocate their resources.

For example, though professional development often fell short of ensuring that kits were universally taught and taught in a particular way, it benefited the sustainability of the program in other ways. It helped to identify teachers with a natural interest in, and enthusiasm for, hands-on science and began to develop in them a more deeply held and more widely shared philosophy that ultimately would serve as a foundation for the programs' endurance. This suggests that as current and future leaders of districtwide science programs develop their strategies, they might consider approaches to implementation that explicitly nurture the development and growth of the philosophy that, as a by-product of kit training, lent critical support for program endurance.

Though all of the factors emerged as important to sustainability, they all are not of equal relevance at every moment or at each level of the program's development. As a program moves through each phase of development, while leaders retain many goals, they also identify new ones that require shifts in planning, implementation, or resource strategy. Further, each additional goal may apply to one or more levels of a school system—meaning , the classroom level, the school level, and the district level—and do so quite differently. For example, leaders' desire to develop a leadership team with an expanded understanding of inquiry teaching may require actions at the school and district level, while their desire to foster the integration of science with reading would require actions at the classroom, school, and district level. Thus, at any given point in the development of a program, program leaders might emphasize a specific factor over others depending on the goal and the appropriate focus level of the system. Together, the phase of development and the program leader's level of orientation determine the factor's importance and priority. See Figure 3 below.

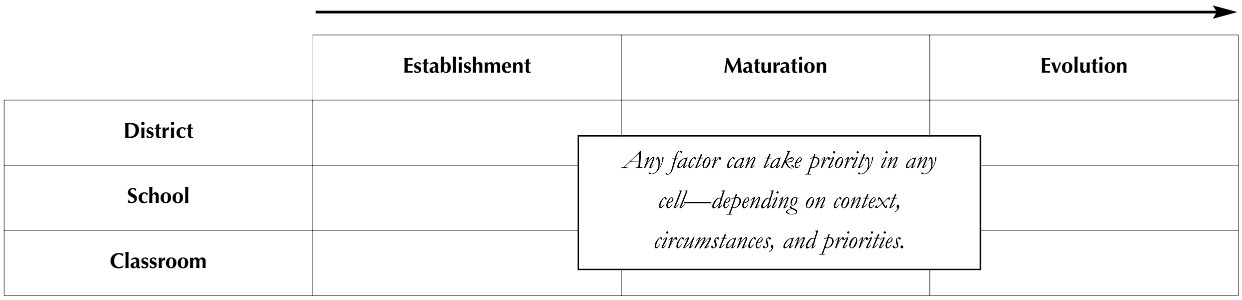


Figure 3

Thus, this study offers readers an illustration of the importance of attending to factors that are not often addressed or even recognized as important to sustainability of a program. It highlights the concrete ways that the programs in this study have done so, albeit most often unintentionally, and offers a starting point for systematically assessing the importance of each, given the particular time and circumstance, and developing strategies to accommodate them.

- *Leaders and supporters of districtwide programs can benefit from defining and considering sustainability through the lens of a long-term time horizon.*

The RSR project's definition of sustainability, discussed in detail earlier in this report (page 19), while acknowledging the factors that pertain to program components highlights the important contributions of the factors that pertain to surrounding conditions and the program as a whole. In particular, it refers to the significance of core beliefs and values (philosophy) and adaptation, and acknowledges the importance of culture and decision making and power as sources of change and pressure:

> *Sustainability: The ability of a program to maintain its core beliefs and values and use them to guide program adaptations to changes and pressures over time.*

This definition of sustainability stresses a shift in understanding from sustainability as program maintenance, in which the elements of the program are preserved over time, to one of adaptation, in which the program elements evolve and adjust.

Given a changed definition of what sustainability is, it is important to recognize that what one should accept as an indication of sustainability changes as well. The evidence that serves to illuminate a program's short-term sustainability is quite different from that which implies long-term sustainability. When educators focus on program elements and see that curricular innovations have been altered or have ceased altogether, or that professional development strategies have succumbed to financial cutbacks, they often conclude that the program has not been sustained. While this may indeed be the case, the RSR definition of sustainability, which addresses sustainability in the long term, asks that educators not reach this conclusion too hastily. Sustainability requires going beyond the establishment phase of a program, which requires vigilant attention to program components, and moving into maturation, which begins to address the more intangible factors that affect sustainability in the long term. Findings from this study suggest that sustainability varies over time and the evidence of it, as the third category of factors suggests, are not always immediately obvious. Factors such as a depth and breadth of shared belief in the value of teaching science and the hands-on approach are critical, but often difficult to discern.

Thus, a look at reform through the lens of this definition of sustainability suggests that it is appropriate to reconsider expectations for the outcomes of program investments. Educators need to recognize that change can be subtle, and it can be latent. And simply because there is no evidence of a "revolution" does not mean that there isn't important evolution. Educators are well-served to reserve judgment about the failure or success of reform until considering all of the ways it may have affected educational practice and interpreting evidence of those changes in light of a long-term definition of sustainability.

- *Leaders and supporters of districtwide programs must increase attention to the quality of their programs with explicit, focused strategies.*

Hand in hand with discussions about how to sustain programs, educators also should engage in a careful and critical look at what is being sustained. The programs in this study were being unevenly implemented and, as such, were not representing the districts' articulated goals for their districtwide programs. Admittedly, many suffered from pressures to bring up student scores in math and language arts, but variable use of the kits also was common during periods when these pressures were much less severe. Even programs at the height of their renown were not being as thoroughly implemented as their reputations would have suggested.

Moreover, just as implementation of a program is uneven, so is the classroom instruction of those teachers who have been using the materials. Despite varying amounts of professional development ranging from several years of saturation following a large NSF grant to sparse and intermittent kit training, there has been little difference in the collective instruction observed from place to place. In each district, instruction is highly variable and, of more concern, its quality is unknown to program leaders. Still, these programs are sustained, often by virtue of the common perception that they are strong, valuable, and having a positive impact.

The issue of program quality is of obvious critical importance to all stakeholders but faces obstacles that prevent leaders from both assessing its value as well as improving it over time. These obstacles, which principally grow out of the lack of authority that provides leaders with access to classrooms and lack of capacity that enables them to collect data and make use of it, are discussed in detail in the section on quality (p. 89). The implications of these obstacles are discussed below with reflections for superintendents, funders, and other policy makers to consider not only for hands-on science programs, but for reforms across subject areas and throughout the education system.

Evidence is essential and beyond the reach of program leaders: Sustaining a program of high quality requires evidence of its impact and its status. The inability of leaders to gather such data is stunning in its absence and chilling in its implications. Data on and illustrations of improved student outcomes can contribute to the value that districts and the public place on hands-on programs, particularly in today's environment of accountability. Lack of evidence of student outcomes, as well as evidence that is not aligned with the goals and intent of the program, leave it vulnerable to being misunderstood and undervalued. In the same vein, without knowing how students are progressing, it is impossible for leaders to know how to direct program improvement efforts.

There is a role for district administrations in building capacity at the district level to address the need for evidence of and data on program impact and effectiveness. The ability to aggregate appropriate classroom data within and across schools would help program and school leaders assemble and interpret the evidence they need to build understanding of and support for the program. The same can be said for the need to access classrooms in order to understand the impact of professional development and program implementation. The discussion of the barriers to assessing implementation points out the challenges principals face in fulfilling this responsibility and the barriers that school-level teacher leaders face in assuming that role with their colleagues. District administrators need to give attention to redesigning roles and considering the distribution of authority so that this function can be fulfilled.

High teacher and principal turnover locks leaders into the cycle of continuous re-establishment and limits their ability to attend to quality issues in the long term: For a variety of reasons, many of which are beyond the control of the school district leadership, teachers enter and leave districts quickly. In doing so, they take with them the investments districts and funders have made in their recruitment, hiring, orientation, and early training.

The costs of high turnover are apparent in all subject areas, but perhaps, more so, in hands-on science programs where professional development is such a critical component.

Districts characterized by a high degree of stability are far more able to advance from the establishment phase into maturation and evolution than districts where teachers and administrators come and go through a revolving door. While they too struggle with questions about the quality of that professional development and its impact on classroom instruction and students, they are better equipped to develop strategies for addressing quality concerns.

If educators accept the premise that professional development is linked to quality of instruction and program implementation, they must recognize the challenge of teacher turnover and account for it if sustained programs are to offer high quality instruction that promotes student learning. This might include redirecting the goals of investments to target building capacity at the district level while designing more cost effective strategies for early, introductory-level training. It might also include directing resources to address retention incentives directly. Regardless of the approach, district and program leaders can not avoid the need to address the threat to stability and lost investments posed by high teacher and principal turnover rates.

IMPLICATIONS FOR THE FUTURE

This study makes the evolutionary nature of reform programs as well as the patterns of disturbances that they endure, explicit. The shocks and pressures that influence programs' sustainability, such as a change in a district's financial status, a shift in public demand for accountability, or decentralization are standard fare and, in response, all districts experience ebbs and flows in the strengths and capacities of their programs over time. Some programs have waxed and waned dramatically, but history clearly shows that all programs, regardless of their age or apparent stability, are vulnerable to shocks and pressures, the majority of which are beyond the control of the program leaders.

In their book, *Tinkering Toward Utopia*[11], Tyack and Cuban observe that the cyclical public debates about education and education reform are driven by "an inevitable result of conflicts of values and interests built into a democratic system of school governance and reflecting changing climates of public opinion." From this source of public opinion come the demands that drive the actions of school districts in their communities, which are further defined by their surrounding conditions—their culture, the distribution of decision making and power, and the understanding of the notion of science for all.

The histories of the nine districts in this study support Tyack and Cuban's notion, and confirm that shifts in societal values and public opinion are pre-

---

[11] Tyack, D. & Cuban, L. (1995). Cambridge, MA: Harvard University Press, p. 41.

dictable, unrelenting, and powerful in their effects on hands-on science programs. Thus, given society's propensity to debate the value of and need for reform efforts and even specific approaches to instruction, any expectation that a sustained program will become immune from these challenges is misguided. Sustained programs are noteworthy, not because they have eliminated threats, but because they survive in the face of them.

Changes in public opinion will never cease. Nor should one expect a resolution to the debates among reformers and practitioners as to the importance of teaching science and the best methods for doing so. The predictable presence of these unpredictable influences raises questions about the future of the programs in this study and others like them. Many have made investments in these programs, and yet they appear to remain vulnerable. This is certainly the case when considered at the close range of program leaders and other supporters, and particularly, at times, like the present, where societal pressure to use "quick" measures of student progress is reverberating throughout the country. Philip Morrison's assertion that "We are not disturbed by slowness, for what goes slow can run deep." faces serious challenges in these times.

This study found that sustained programs withstand these potential threats with resilience that lay in strengths not easily seen. They were in places where no one had looked—meaning, in the more subtle factors of adaptation, perception, philosophy, and critical mass—and were apparent only after the passing of time. Understanding sustainability from the perspective of history and these more subtle factors does not guarantee better outcomes for hands-on programs. But, it does argue that, if leaders attend explicitly to what were previously unrecognized program supports, as debates arise about the way science should be taught and the worth of hands-on programs, their value will be explicitly and thoroughly presented. Likewise, when more hospitable times return, programs will be better equipped to advance further, with greater confidence in their awareness of the gains they have made.