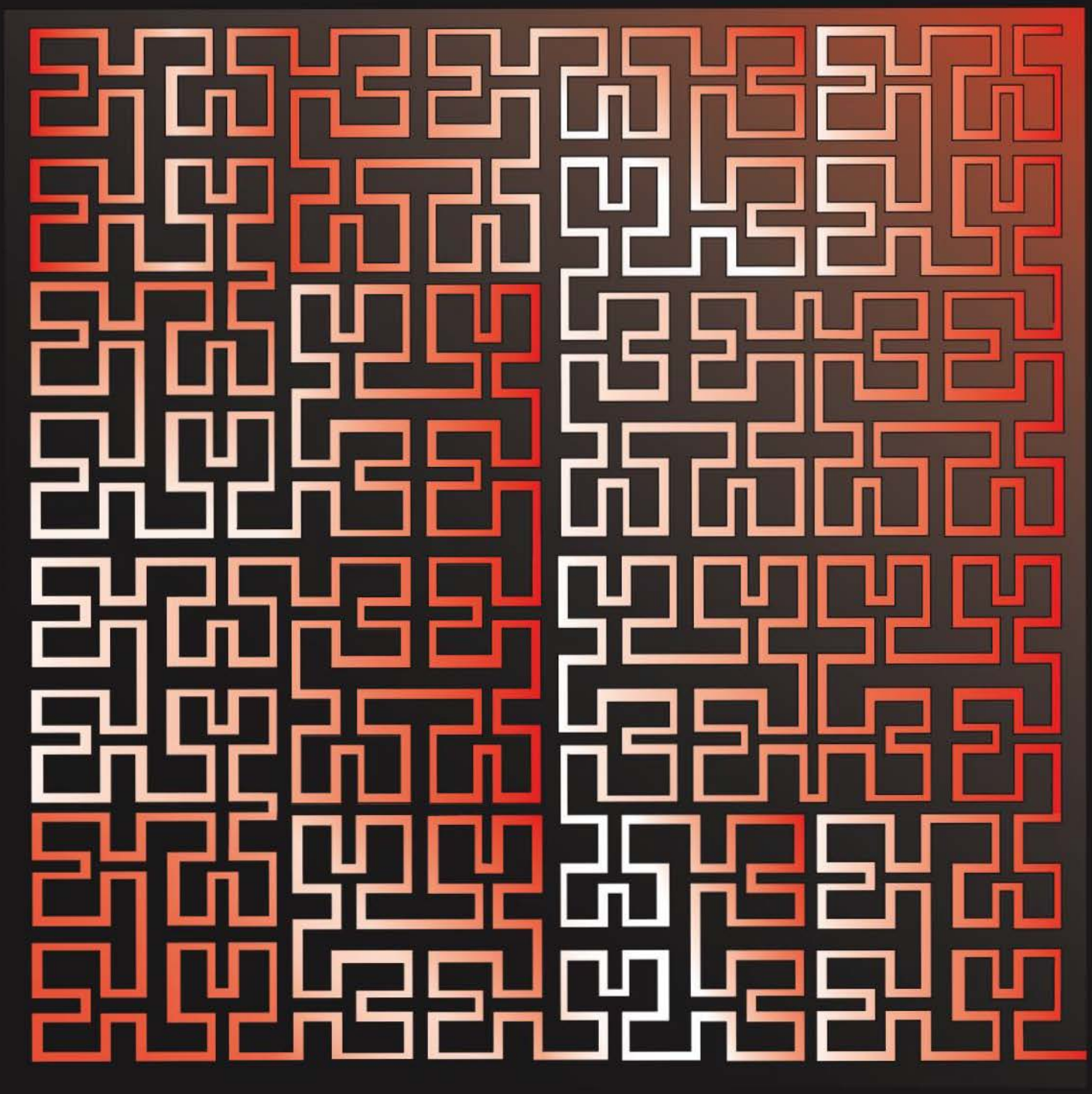
SCIENCE EDUCATOR

THE NATIONAL SCIENCE EDUCATION LEADERSHIP ASSOCIATION JOURNAL



Spring 2010 • Volume 19, Number 1

EDITOR: Jack Rhoton
East Tennessee State University

Science Educator

Spring 2010 Vol. 19 No. 1

Table of Contents

Manuscripts: Submit manuscripts to the Editor, Brenda Wohnowski, North Texas University, Teacher Education and Administration, 1155 Union Circle #310740, Matthews Hall 206R, 1300 W. Highland Street, Denton, TX 76203-5017. Refer to the information for authors elsewhere in this journal. The opinions and statements published are the responsibility of the authors, and such opinions and statements do not necessarily represent the policies of NSELA. Annual membership dues for NSELA are $75.00, $30 individual journal-only. Non-members will be assessed a charge of $10.00 per issue. For subscription orders and customer service, call (423) 439-7589.

The *Science Educator* (ISSN 1094-3277 Copyrighted 2010 by the National Science Education Leadership Association (NSELA) is published biannually with provisions for quarterly publication in the future. Printed at the University Press, East Tennessee State University, Johnson City, TN 37614, the journal serves as a forum for presentation and discussion of issues pertaining to science education leadership. It is indexed in the ERIC Clearinghouse for Science, Mathematics, and Environmental Education. General inquiries may be directed to: Office Manger, Judy Hamilton <judyhamilton51@gmail.com> or Membership Co-Chair, Beth Snoke Harris <beth@seven-oaks.net>, or Brenda Wohnowski, Editor, North Texas University, Teacher Education and Administration, 1155 Union Circle #310740, Matthews Hall 206R, 1300 W. Highland Street, Denton, TX 76203-5017 [Ph: (940) 565-2920; e-mail: brenda.wohnowski@unt.edu].

Xiaodong Zhang, Joseph McInerney, Joy Frechtling

Engaging STEM Faculty In K-20 Reforms—Implications for Policies and Practices

This article looks at policies and strategies that can be used to promote partnerships involving university science, technology, engineering, and mathematics (STEM) faculty and K-12 teachers, as well as the nature of such collaboration.

Introduction

The Math and Science Partnership (MSP) program at the National Science Foundation (NSF) is a major national research and development effort that supports innovative partnerships among institutions of higher education (IHEs), local K-12 school systems, and their supporting partners in order to improve K-12 student achievement in mathematics and science.[1] Deep engagement of science, technology, engineering, and mathematics (STEM) disciplinary faculty is a hallmark of the MSP program. The program posits that disciplinary faculty hold the content knowledge that K-12 teachers need and that, if faculty are substantially involved, teachers' disciplinary knowledge will be strengthened, resulting in improved student achievement.

Many reforms stress partnerships among institutions of higher education, K-12 schools and districts, and community-based organizations and businesses (Abbott et al., 1992). The MSP program and the U.S. Department of Education's Teacher Quality Enhancement effort have the explicit goal of forming partnerships between K-12 districts and IHEs in order to create innovative solutions to persistent instructional problems and lead to improvement in both K-12 schools and IHEs. Educational partnerships between universities and public schools are not new. There are, however, three reasons for the current interest surrounding such partnerships. First, the politics of education reform have created the need for at least symbolic association among educational stakeholders. Second, increased accountability for student achievement, coupled with the need for better-prepared teachers, has placed pressures on public schools and IHEs to collaborate. Finally, K-12 schools and IHEs face similar problems, such as public criticism, lack of sufficient funding, limited public support or respect, low salaries, and faculty shortages (Sirotnik & Goodlad, 1988). According to Teitel (1999), common interests have brought together a strong convergence on four goals: improvement of student learning, preparation of educators, professional development of educators, and research and inquiry into improving practices.

Although partnerships are easy to extol, they are difficult to achieve.

Although partnerships are easy to extol, they are difficult to achieve. One of the most prominent reasons for this difficulty is the institutional reward structure, which puts different emphasis on research, teaching, and service (Boyer, 1990). According to Diamond (1999), an appropriate and effective tenure and promotion system should be aligned with the institution's mission statement; be sensitive to differences among the disciplines and individuals; include appropriate, fair, and workable assessment; and recognize that action takes place at the departmental level in which the most specificity in documentation is required. Although many IHEs' mission statements recognize teaching, research, and service, there is often a mismatch

in reality between the mission of an institution and the priorities described for the tenure and promotion systems. As Boyer noted in 1990, "almost all colleges pay lip service to the trilogy of teaching, research and service, but when it comes to making judgments about professional performance, the three rarely are assigned equal merit" (p.15).

A 1996 survey of 50,000 faculty, chairs, deans, and administrators at research universities (Gray, Diamond, & Adam) showed that respondents often considered the balance between research and teaching on their campus to be inappropriate. A more recent national survey (Alshare, Wenger, & Miller, 2007) found that deans at teaching universities, on average, assigned percentages of 47/43/10 to teaching, research, and service activities for promotion and 48/42/10 for tenure decisions. In contrast, deans at research institutions assigned percentages of 59/33/8 to research, teaching, and service for tenure and 57/32/11 for promotion. The difference between teaching and research universities lies largely in the relative weights on teaching and research, but it is clear that service is a distant third in both cases.

Tenure and promotion is a powerful motivator to faculty (Colebeck, 1994). However, the increased prominence of the research enterprise and lack of rewards for public service have contributed to the socialization of faculty away from public service, even at institutions with strong service traditions (Jaeger & Thornton, 2008). As a result, faculty have been forced to exhibit market-like behaviors to secure competitive funds from government grants or the private sector and ignore teaching and service.

Faculty and administrators are often prisoners of their own thinking, firmly holding values about faculty roles, scholarship, and institutional identity shaped by the current reward system that promotes a "publish or perish" culture.

Faculty and administrators are often prisoners of their own thinking, firmly holding values about faculty roles, scholarship, and institutional identity shaped by the current reward system that promotes a "publish or perish" culture (Senge, 1990, p.27). A 2005 survey of 729 chief academic officers found that two-thirds of respondents believed that faculty graduate school training and socialization toward traditional forms of scholarship served as a barrier to encouraging multiple forms of scholarship (O'Meara, 2005). Another reason that service is devalued may relate to a lack of means to assess quality in public service. In research, the universal language of exceptionality is the number of publications in top field journals, an easily countable and recognizable measure. McDowell (2001) claimed that public service and teaching were often overlooked in promotion, because proper evaluation of achievements in these areas was more difficult than mere counting.

Although the current tenure and promotion system seems deeply entrenched, it has not always been this way. Boyer (1990) pointed out that the missions of universities have changed throughout the years—moving from teaching, to service, and then to research, in response to shifting priorities both within the academy and beyond. However, at the very time the mission of American higher education was expanding after World War II, the faculty reward system was narrowing to its current status. Fortunately, some colleges and universities have attempted to change the existing tenure system (Chait, 1998). Two approaches have been successful in encouraging faculty engagement in K-20 partnerships. One approach is to elevate the status of service, which is often how faculty involvement is defined. Another approach is to redefine scholarship to include teaching, discovery, integration, and application of knowledge.

In addition to changing the tenure and promotion system, Boyer (1990) argued that universities should also create flexible and varied career paths for professors throughout a lifetime in order to counter burnout or stagnation. He observed that late-career professors may experience a peak in status and recognition, and that it is at this time that demands for their service from outside their institution often grow. The argument for career flexibility relates to variations in the disciplines, since patterns of productivity vary from field to field. STEM faculty, for example, are often most productive in their youngest years. In the STEM fields, it is common to devote most of one's early career to specialized research and then turn to integrative questions. At this later stage, faculty might take time to read in other fields, write interpretive essays or textbooks, or collaborate with a colleague on another campus. Still later, the faculty member may focus on an applied project.

Drawing on a larger study that examines the effects of STEM faculty engagement in MSP, this article specifically looks at the tenure and promotion policies in a sample of IHEs involved in MSP. Recognizing that tenure and promotion policies may be slow to change, we also examine strategies that were used by selected MSP projects to engage STEM faculty in K-20 educational reforms in the absence of major policy changes at the university level.

Methodology

The findings are based on case studies of eight MSP projects. The eight case study projects were selected from a pool of 48 projects, primarily because, based on their proposals, these projects were expected to include high levels of STEM faculty participation. Two of the eight projects focused on mathematics, three on science, and three on both mathematics and science (Table 1). Among the lead institutions, four are classified under the Carnegie classification system as Research University (very high research activity), one as Research University (high research activity), one as Doctoral/Research University, and two as Master's College or University (larger program). Six of the IHEs are public and two are private. Geographically, they are located in the East, Midwest, South, and West. The number of IHEs within a partnership varies from 1 to 10. The number of K-12 districts ranged from 2 to 29, with an average of 10.

Although the current tenure and promotion system seems deeply entrenched, it has not always been this way.

In addition to document reviews of tenure and promotion policies, we conducted annual site visits to the eight projects in order to describe faculty engagement over four years and to identify changes that occurred. Site visits often included interviews (with project leadership, STEM faculty members, department chairs, in-service teacher leaders and teachers, principals, and district content/curriculum specialists) and classroom observations of STEM faculty and K-12 teachers with whom STEM faculty have worked. The interview questions were linked to respective research questions that reflected both the roles of the respondents and the maturity of the project. The semi-structured, open-ended question format allowed for additional questions or probes to be used as deemed necessary

Table 1: Characteristics of the case study sample

Aspect	P1	P2	P3	P4	P5	P6	P7	P8
Content focus	**M**	**S**	**M/S**	**M/S**	**M/S**	**S**	**M**	**S**
Institution type of lead partner								
Carnegie classification								
Research University (very high research activity)			X		X		X	X
Research University (high research activity)	X							
Doctoral/Research University				X				
Master's College or University		X				X		
Ownership								
Public	X	X	X		X	X	X	
Private				X				X
Location								
East				X		X		X
South	X							
Midwest					X			
West		X	X				X	
Total number of IHE partners	2	5	1	2	5	4	10	1
Total number of K-12 district partners	15	29	3	10	8	2	10	17

P = Project; M=Mathematics, S=Science.
Source: MSP MIS; case studies.

Although the majority of the participating STEM faculty were highly motivated, they still needed additional incentives to sustain a high level of motivation; that is, self-motivation is not enough, especially when projects require extensive multi-year involvement from the faculty.

by the interviewer. We conducted observations, including those of non-participants, in classrooms and at a variety of meetings. The annual site visits were conducted by teams of two that consisted of a researcher and a STEM disciplinary faculty member from a non-MSP university. The inclusion of a STEM faculty member as co-site visitor helped to establish a rapport between respondents and researchers, and it also provided insight that allowed the data to be interpreted in a more culturally sensitive way.

The literature guided some of the coding and analysis of interview notes and documents (Patton, 1990), but codes emerged primarily from the data. The data analysis followed the process of 1) developing preliminary coding categories from the research questions and confirming or modifying those categories as information was gathered, 2) reducing the data to manageable chunks of information for identifying themes or patterns of response, and 3) drawing conclusions by comparing within-case and across-case themes and patterns (Miles & Huberman, 1994). Essentially, we used two forms of triangulation. Within each project, evidence was triangulated from interviews and observations. Across projects, evidence was compared and contrasted in the context of each project.

Results

The number of STEM faculty involved in the eight case study projects varied considerably, from 8 to 50 with an average of 22 per project (Table 2). The majority of the participants were tenured or tenure-track faculty. Faculty participation usually involved two to eight weeks over the summer, depending on the length of the summer institutes. For projects that required commitment during the school year, the extent of involvement varied markedly—from two days a month to 50 percent of the participants' time. STEM faculty devoted considerable time in the areas of in-service and pre-service teacher training, curriculum development, project management, and research.

Although the majority of the participating STEM faculty were highly motivated, they still needed additional incentives to sustain a high level of motivation; that is, self-motivation is not enough, especially when projects require extensive multi-year involvement from the faculty. In fact, the issue of incentives may be even more critical to further expansion of STEM faculty engagement, especially as the current IHE reward structure and tenure policies are not conducive to MSP-like activities.

IHE Tenure and Promotion Policies Related to MSP Activities

Tenure and promotion policies were among the main foci of our investigation, because they are often considered one of the biggest hurdles to creating K-20 partnerships. Our focus was not only on how such policies were articulated at the university (macro) level but also how they were implemented at the department (mezzo) level and how they were perceived by the STEM faculty themselves (micro-level). We found that research and sometimes teaching were the principal paths to

Table 2: Extent of STEM faculty involvement in MSP in the case study projects

Aspect	P1	P2	P3	P4	P5	P6	P7	P8
Number of STEM faculty involved in development/ delivery of MSP activities	8	29	50	14	36	10	21	14
Percent tenured or tenure track	100%	72%	78%	57%	64%	100%	81%	57%
Amount of involvement								
Percent more than 40 hours	100	95	30	100	100	83	92	88
Percent more than 160 hours	38	74	14	64	35	83	77	31

P = Project.
Source: MSP MIS IHE Institution Survey 2005-06; case studies.

tenure and promotion. Considered to be a distant third path, service or outreach, sometimes referred to as "other contributions," was used to define MSP involvement in most cases. In order to survive in a highly competitive IHE environment, large research universities sought to maintain their edge, and smaller colleges strived to become research universities. Nevertheless, service and/or outreach were not completely ignored. Of the eight lead IHEs, the policies of seven contained language recognizing service and/or outreach and five recognized publications that pertain to teaching and learning.

It was widely acknowledged that tenure and promotion policies were key to encouraging STEM faculty participation, and one may hope that this realization would lead the universities to review and modify policies so as to create an environment more conducive to faculty engagement. Unfortunately, our eight case policy projects showed little change at the university level in these areas over the last four years. Most MSP projects were not designed to change IHE tenure and promotion policies and practices. Only one of the eight case studies articulated influencing the reward structure for IHE faculty as a goal. That project planned to convene a 15-member Council for Math, Science and Education in order to address the larger issues that plague education reform, systemic problems, and specific issues arising from the project's operation. Potential issues included reward structures for IHE faculty and K-12 teachers, institutional change, and sustainability. However, this activity never really started. In fact, of all 48 MSP projects in the first three cohorts, only one project has made significant change in the university's tenure and promotion policies.

Tenure and promotion policies were among the main foci of our investigation, because they are often considered one of the biggest hurdles to creating K-20 partnerships.

Despite official recognition at the university level, university policies were often implemented differently at the department level where "the change is really taking place." In at least two instances, department policy statements noted that the department generally avoids major service demands on untenured faculty and that leadership in outreach/service is not part of the criteria for tenure. However, sections of the policy statements on appointment to full professor mentioned the need to demonstrate significant accomplishments within the department, university, and professional societies, as well as outreach to the community, including civic duties related to mathematics and science education. It was not surprising that most of the faculty members participating in the MSP project were tenured, so the younger, less established ones did not have to "sacrifice" time that could otherwise be spent conducting research. In fact, one MSP project director was told that tenure-track faculty were "off limits." Many respondents observed the same basic pattern. Junior faculty are required to focus on research and teaching first. Then, once tenure is achieved, the balance of responsibilities may change, and faculty may begin to either focus entirely on research or become engaged in teaching or service. As a result, tenured faculty have much more freedom to decide how they allocate their time and resources.

We did see some incremental changes in practices at department levels. For example, one lecturer worked with the MSP from the beginning. She had her five-year review and was certain that her work with the project was responsible for a "larger than normal" salary increase, because the review committee highlighted that work in their written report. In another project, two participating faculty received tenure and were promoted to the rank of associate professor. Both credited their involvement with the project as having played a positive role in the promotion and tenure decisions. However, we also heard stories that outstanding MSP participants were denied promotions due to a lack of research.

Tenure and promotion policy and practice changes in the departments were more likely to take place when department chairs were involved in MSP. One PI, who was also the department chair at a research university, said "as long as I am chair, it will play positively in terms of tenure and review." That was not true, however, with all projects. Another PI who has been the chair of the department said that there had never been any intention to change the promotion and tenure criteria in his department to recognize service more favorably.

In addition, efforts were made to redefine MSP activities in terms of research or teaching. For example, up to the second year of our study,

MSP or STEM education-related research was identified only as an area of STEM faculty activity in one of the eight case study projects. Our latest round of visits found that at least five projects have faculty conducting STEM education research. The actual implementation can be layered and complex, as shown below.

- In his evaluation of faculty members, a math department chair who has been directly involved in the project from the beginning defined MSP as "multi-disciplinary and collaborative work." The college considered it in tenure and promotion decisions, because it combined multiple components. In addition to outreach, research in mathematics education was counted as "application of math" similar to its applications in engineering or statistics. However, few participating faculty members had yet based their research agenda on their MSP activities. The chair added "if they do that, I will argue for it … I'd prefer to count it as research contribution, because it is more highly regarded in the community."
- In another case, a department chair was not directly involved in the project but had provided moral support and space. Although we heard that research on the scholarship of teaching was recognized at the university, the chair's comment showed that the influence of those activities on advancement decisions was still ambiguous. He stated, "It is difficult. Although service is valued, research weighs heavily. We have not totally figured it out. It is tricky to contextualize. People are always suspicious about publications, and it has to be on a case-by-base basis with more justification." The chair continued, "There is a lot of sympathy. The ongoing focus has to do with the proportion of academic involvement in the outputs. The administration is reasonably receptive."

It was widely acknowledged that tenure and promotion policies were key to encouraging STEM faculty participation, and one may hope that this realization would lead the universities to review and modify policies so as to create an environment more conducive to faculty engagement.

Another area of change was in hiring practices through the creation of tenure-track positions for teaching faculty or STEM education researchers in the STEM departments. For example, a department chair noted that there had been a new faculty slot added for STEM education. "Right now, there is not a critical mass, but I am not surprised to see it develop into a new sub-program. It will fit in the department nicely." Similarly, for another project, the department hired a tenure-track STEM educator who will spend 80 percent of his/her time in disciplinary research and 20 percent in education research. The chair noted, "This would have been laughed away five years ago," but he felt that people were starting to understand. For the third project, two STEM education faculty were hired at the lead institutions, which may be the result of MSP or increased awareness within departments of the need for STEM education insights.

Several departments in another project have made hiring discipline/education faculty—e.g., doctorates in physics education and mathematics education—a priority. In one instance, a department chair asked his faculty to make a choice between beginning a search for a biosciences education faculty member or a marine biologist. The educator position was chosen. Departments that elected to move in this direction have found it to be an "exceedingly difficult and competitive" undertaking due to both the lack of people with these credentials and the increasing number of IHEs attempting to attract those who are available.

The ultimate goal of these policies and practices is to influence faculty so that they will be more attracted to service or to the scholarship of teaching and/or engagement. When asked whether they would be rewarded at their institutions for participating in an MSP-type activity, a minority of the STEM faculty thought that participation was viewed positively; most felt it was either tolerated or ignored. Some felt that the most that could be hoped for was for deans and department chairs to broadcast a message indicating "there is no reward for doing this, but it is okay for you to do it." Most faculty members believed that teaching and service would never make up for the lack of research. However, it also appeared the distinction was not necessarily set in stone. One faculty member told us,

"It would be up to me to characterize it and present it to the university." In his case, MSP work, especially the professional development piece, was defined by the department as service, and pre-service teaching and curriculum design fell under teaching and curriculum development. However, if the MSP work came out in peer-reviewed journals, it would be classified as scholarship.

As we kept probing STEM faculty about changes in university tenure and promotion systems, we found little movement in their views. One STEM faculty member observed that "any consideration of coupling the three areas (research, teaching, and service) as equal is moving slower than a glacier." Changes regarding institutional reward and tenure policies will continue to be slow and can be controversial, even among the participating STEM faculty. As one noted, "I am ambivalent about it. Achieving tenure through outreach will create different attitudes. I don't think outreach should be an easy way to get tenure." One co-PI said that IHE policies have not changed, but a foothold has been established in the university for thinking about MSP-style involvement. "Policies, no; mind sets—there has been a change."

Project Strategies to Engage STEM Faculty

Although tenure and promotion policies are critical to engaging STEM faculty, most MSP projects were not specifically designed to tackle that issue directly. In addition, changes in university policies can be slow to take place. Nevertheless, there were a number of effective strategies projects used to increase STEM faculty engagement in the absence of major changes in institutional tenure and promotion policies. Motivation for STEM faculty to become engaged in a multi-year project like MSP appeared to hinge on two necessary and entwined conditions. The first condition was extrinsic and clear. Projects needed to provide adequate course release and/or stipends for participating STEM faculty. The second condition for STEM faculty engagement was an intellectual connection. That is, the project needed to make the case for the need for substantive STEM faculty work with K-12 teachers. This was an intrinsic and, perhaps, underestimated condition. Using evidence from the case studies, we synthesized these strategies into two categories: providing extrinsic incentives and providing intrinsic incentives.

Extrinsic incentives

All eight projects offered stipends, and five provided release time as extrinsic incentives. These incentives were established at the beginning of the projects and have remained consistent over the period of MSP. If the STEM faculty were involved in summer institutes, the stipends often were for one or two months. However, one PI was adamant that offering support for a minimum of three months would make it easier to secure faculty commitment. For one project, stipends were larger during the first project year, when courses were being developed, and smaller for other years, when only modifications and adjustments were needed.

Involvement during the school year was normally compensated by release time and/or stipends. Of the three projects that did not provide release time, two were projects whose primary activities took place in the summer. In the third project, faculty participation was originally planned to occur in the summer. However, many teams decided to conduct at least some of the activities during the school year, and the incentive scheme was not revised, so faculty continued to be reimbursed with stipends but not release time.

Providing stipends during the school year can be complicated. University policies on faculty consulting or "overload" may require a considerable amount of paperwork or restrict the amount of compensation that a faculty member may receive in the form of stipends. The rate for stipends may also vary. Some faculty members were willing to take a few hundred dollars for their contributions, whereas others requested that federal government consulting rates be used as a reference point.

Compared to stipends, release time was more difficult to get, especially in IHEs that were more teaching oriented. One course release per term seemed to be the norm. For one project, MSP teaching counted as part of the teaching load; in others, release time had to be negotiated. For example, a department chair had to make a strong case with the administration to arrange for course buy-outs for faculty, because, in his institution, release time was normally possible only in research-related situations. Practices often varied within a project. It was often the case that, while faculty from a lead institution might receive a course buy-out, members from non-lead institutions did not receive any course release.

Although we recognize the importance of extrinsic incentives, we do not ignore the element of altruism. In fact, many participating STEM faculty suggested that they got involved primarily because they were concerned about public education and wanted to serve the local community

and make a difference or because they simply enjoyed teaching. As one PI observed, "It is less about incentive and more about people interest." We came across cases in which faculty members had no idea or did not care about how much they got paid for the involvement. There were also cases in which faculty did not seek course release, because they enjoyed teaching so much. "Teaching teachers is the best part—the reward," one STEM faculty added.

The ultimate goal of these policies and practices is to influence faculty so that they will be more attracted to service or to the scholarship of teaching and/or engagement.

However, most STEM participants acknowledged that both stipends and release time were attractive, and they, naturally, "appreciate the extra money." Some noted that the idea of servicing the community may appeal to the more established faculty, while the monetary incentive was more appealing to the younger, less established, non-tenure-track faculty. For community college faculty who have much heavier teaching loads, stipends and release time seem to be more crucial. In general, faculty felt that their involvement was supported generously by the projects, although one PI argued that there was not enough support (money or release time) for faculty involvement and "the good nature of faculty is always needed to make up for any insufficiency."

Intrinsic incentives

In addition to providing summer stipends and course release, projects employed a number of strategies to appeal to intrinsic motivations. These strategies include selective recruitment of faculty, professional development, effort to promote collaboration, and projects' sensitivity and flexibility to faculty needs.

Selective recruitment

One department chair stressed the importance of combining "money talking" and enlightened self-interest when engaging faculty. It often fell upon the project leadership to actively engage STEM faculty. "The PI needed to beat down doors at the university to get more scientists involved," said one co-PI. Several respondents pointed out the importance of finding STEM faculty with genuine interest in education who were willing to extend themselves rather than to say, "I have all the answers." Engaging in K-20 reforms is not for every STEM faculty member. Some case study projects mentioned the importance of selective recruitment. "There are STEM faculty who I love and respect, but I wouldn't let them near the project," said one MSP project leader. Although it is difficult to generalize, one common element was that many of the STEM faculty in the case studies initially became involved with K-12 schools because they had children in them or their spouses were K-12 teachers.

Many respondents noted that it takes a "certain type of personality" to be effective in K-20 partnerships. An ideal STEM faculty participant often has the following traits: 1) possesses a high-quality disciplinary background and credibility, 2) is a good STEM higher education instructor and interested in how to teach more effectively, 3) has a dedication to changing the lives of students, 4) is open-minded to trying new approaches, 5) is able to deal with people who are coming from different content-level foundations, 6) is willing to work in teams, and, on a lighter note as per one PI, 7) is "in touch with their inner adolescent." A project evaluator summarized that the key to engaging STEM faculty was to use time well, compensate them with money and opportunity to collaborate, and make them feel that their voices are heard. Otherwise, "they will vote with their feet."

While most projects recognized the importance of building partnerships early on, some projects had a steep learning curve throughout the years about the value of providing professional development for STEM faculty and demonstrating sensitivity and flexibility to faculty needs. In general, projects that made a substantial case for reform—that is, they laid the intellectual groundwork for new roles and models of STEM faculty engagement with K-12 teachers early on—reaped the benefits as the project progressed. Project leadership was critical in establishing such groundwork. Meaningful and prolonged STEM faculty engagement hinged on the balance of the two motivating conditions or, as one respondent put it, "the practical piece and the learning piece."

Professional development

MSP has high expectations of STEM faculty as an agent of change. However, just having a PhD in a STEM discipline is insufficient for the task, because doctoral systems are designed to produce researchers, not educators. As one PI, a chemist, bluntly pointed

out, "STEM faculty are typically clueless. They don't understand the content needs of K-12 teachers. They don't know where to start. And once they've gotten started, they don't know where to go."

Professional development for STEM faculty was an area of considerable growth over the years for MSP projects. Almost all case study projects had provided some forms of professional development for their STEM faculty, even though some was less intentional and intensive than others. For example, the training could be periodic (e.g., monthly) meetings or debriefings after workshops during which faculty discussed among themselves and/or with other participants general issues such as course content, methods of presentation, texts, and program requirements, or specific issues on research, curriculum development, and assessment. Other professional development was more systematic and intensive. Two examples of intentional professional development follow:

- One project organized biweekly seminars involving participating STEM faculty, education faculty, and graduate students. In each seminar, participants discussed the books and literature they read about best practices regarding the types of courses that need to be offered in teacher preparation programs and in-service professional development programs and the ways in which these courses should be taught. According to the PI, the biweekly seminar was the key to success, because participants forged common language, knowledge, and experience with each other. A senior STEM faculty said, "This is the methods class that I've never had before."

Although tenure and promotion policies are critical to engaging STEM faculty, most MSP projects were not specifically designed to tackle that issue directly.

- Another project devoted an entire summer to providing professional development for STEM faculty members and teacher leaders on pedagogy and exemplary middle school curriculum materials *before* the teams were assigned to work with schools. Faculty contributed by assessing the curricula with regard to its effectiveness in identifying the kind of thinking needed in college, addressing the problems that students had in moving from the concrete to the abstract, and improving the scientific sophistication of the lead teachers. They learned from teacher leaders about school contexts, student diversity, and state curriculum standards and assessments.

Seriously engaged faculty did not view their involvement traditionally in terms of outreach or service roles but instead spoke of what they gained intellectually and professionally from participation. One project director reported that faculty "want professional development for themselves," because they "want to learn these things." Some examples of this intellectual engagement are as follows: frequent mentions of foundational MSP readings such as *How People Learn*; references to the IHE faculty professional development they experienced prior to working with teachers; long-term collaboration with K-12 science and mathematics teachers and teacher leaders; insights from the field of science and mathematics education research; participation in a Lesson Study or the examination of student work with an emphasis on student understanding; change in their own disciplinary teaching as a result of the MSP influence; discussions about student preconceptions, cognitive load, and questioning strategies; mention of a forthcoming publication in a STEM education journal; and presentations at a National Association of Research in Science Teaching meeting.

Promoting collaborations

Partnership formation plays a critical role in MSP program delivery and in fostering a climate to support institutional changes. The idea of partnership includes collaboration not only among IHEs and K-12 school districts, but also among the STEM faculty and other project participants. Importantly, STEM faculty typically work in an independent fashion and may not be comfortable in a collaborative environment in which people come from different backgrounds and have varying levels of content knowledge. Consequently, establishing and maintaining a true collaborative environment, especially setting up the framework, is critical to project success as well as to STEM faculty engagement.

In six of the eight cases, STEM faculty worked in teams with teacher leaders and/or education faculty, often in activities such as providing summer institutes for in-service teachers.

Essentially, many of these projects were built on the "co-learner model," although STEM faculty roles varied from leading to supportive. Projects specified the type of roles to be filled by participants in some cases, while in others it was left entirely for the team to decide. Four examples follow.

- For one project, collaboration among participants was inherent in its operational model, which is known as teacher research teams. The team was composed of college disciplinary and education faculty, high school teachers, and undergraduate and high school student tutors. The hypothesis was that teachers will improve teaching skills on the job by working in teams with supportive instructional staff and content experts to conduct summer camps for high school students who failed the state exams. A number of features were built into the system to encourage cooperation, at least in theory. For example, faculty and high school teachers spent a week working as a team to prepare curricula for the summer program. They delivered the instruction as a team. At the end of each day, each team spent one or two hours debriefing and reflecting about the day.
- A second project was decentralized in 10 school-based teams. Each team included two IHE faculty, most of whom were STEM faculty. Other team members were teacher leaders, a principal, and a guidance counselor or school social worker. Working as colleagues one to two days per month, IHE faculty and K-12 personnel tackled school issues in STEM education and learning.
- A third project had a very unusual arrangement. During the three-week summer institutes, faculty and K-12 teachers were required to be in residence. This aspect of the program was credited with having created a bonding and a professional learning community that could not have been achieved through other means.
- Another project brought together faculty from all levels—grade school, middle school, high school, community college, and university—by focusing on a dialogue about a particular mathematics or science concept. Although most participants felt that they were able to connect informally with STEM faculty and develop some valuable relationships, the activity was discontinued due to concerns that the episodic and "short-term" nature of the activity was not likely to influence teachers and students.

Bringing people together is one thing, but making it work is another. One PI pointed out that "willingness to work as a team is the toughest part." Although many of the collaborators exhibited collegiality and camaraderie, some teams encountered problems. In some cases, the PIs had to change teams in order to resolve issues related to personality conflicts. In other cases, these problems were not resolved, and the affected members withdrew from the program. Fortunately, we observed an increasing ease in communications between the STEM faculty and other members across projects over time.

While most projects recognized the importance of building partnerships early on, some projects had a steep learning curve throughout the years about the value of providing professional development for STEM faculty and demonstrating sensitivity and flexibility to faculty needs.

Sensitivity and flexibility to faculty needs

Projects have become increasingly aware that they must be sensitive to the priorities and needs of STEM faculty. Perhaps the number one issue is time. STEM faculty are often constrained by multiple and sometimes competing demands. For some projects, the fact that the majority of activities occurred in the summer reflected projects' intentions that "research does not need to take a major hit," although some faculty regarded summer as the optimal time to do research. One project's experience is particularly illustrative. In the first year, STEM faculty expressed a concern about being stretched too thin by multiple responsibilities and time demands. The project changed strategy by requiring intense STEM faculty involvement only during the year in which their content area was featured in the summer institute.

Publication is another need of STEM faculty, and publications about

disciplinary education require support and mentorship. In an attempt to capitalize on the experience gained from MSP work, one project began to hold seminars about disciplinary journals that target STEM pedagogical research. With impetus from the project, the university established a research network promoting pedagogical research in STEM disciplines. Another project offered a writing seminar that provided faculty an opportunity to chronicle their research findings and document the curricula they had developed.

Flexibility is also an important consideration. One project involved both mathematics and science faculty. The science faculty worked with education faculty on the instructional team and sought an increased variety of pedagogical strategies, such as differentiated learning for both pre-service and in-service teacher training. The mathematics faculty, however, chose to focus on other goals, because they felt that such an arrangement would make it hard to recruit content-focused colleagues. "We are being tapped by the MSP for what we know—content, and not for what we don't know—school pedagogy," one faculty member told us. The project did not try to impose one approach over another. Instead, difference between these two approaches appeared to be a "benign bifurcation."

For another project, STEM faculty were frustrated that they could not integrate their experience into their professional lives after two years of MSP engagement. The project redesigned its faculty involvement plan and left it up to faculty to choose how they wished to participate. As a result, some focused on pedagogy, some on content, and others on research. Only two faculty members stayed with the original concept of involvement. In general, STEM faculty took roles in the schools that best corresponded to their area of expertise, interests, and comfort levels.

Conclusions

MSP projects employed a number of effective practices to support faculty involvement. Extrinsic incentives were universally used, whereas intrinsic incentives were sometimes underestimated.

- At the project level, both extrinsic and intrinsic incentives needed to be created. The former often involved providing stipends and release time for faculty members, and the latter often included providing professional development to faculty in order to enhance their understanding of K-12 perspectives and pedagogical issues, as well as building collaborations among participants and demonstrating sensitivity and flexibility to faculty needs.
- Extrinsic incentives were well understood, as evidenced by the finding that all of the case study projects offered stipends and five provided release time. These incentives were established at the beginning of the projects and remained consistent over time. The intrinsic piece—making the case and creating the intellectual connection for substantive STEM faculty work with K-12 teachers—was often underestimated. Although most projects recognized the importance of building collaborations early on, many projects had a steep learning curve about the value of providing professional development for STEM faculty, as well as demonstrating sensitivity and flexibility to their needs.
- Traditional tenure and promotion structures and faculty perceptions about the status associated with different types of engagement were considered major barriers for faculty involvement in most MSP-like endeavors. Although the majority of the IHEs recognized service or outreach, these types of activities were generally considered to be a distant third in priority after research and teaching, and this presented a major roadblock to involving faculty from the STEM disciplines. As a result, the majority of participating STEM faculty already were tenured. Although tenure and promotion policies were critical to STEM faculty engagement, most MSP projects were not specifically designed to tackle those issues.
- Changing tenure and promotion policies was a slow process. We found that small steps had been made toward either elevating the status of outreach/service directly or redefining MSP activities in terms of research or teaching. However, faculty perceptions about tenure and the reward system remained the same.

Since most of the IHEs in the MSP are research universities and the participating STEM faculty are either tenured or on a tenure track, these findings are limited to these IHEs and faculty. However, the potential for STEM faculty involvement is much

larger than the level of involvement that is currently being realized. It may be worth considering the possibility that MSP would be more successful in recruiting faculty from IHEs that are less research-oriented, because STEM faculty from research universities are less likely to be interested in K-20 reforms. In contrast, STEM faculty from liberal arts colleges and regional state universities may be more inclined to engage in such efforts, because teaching and learning forms a more integral part of their mission, and, consequently, faculty from these institutions are more likely to reach out to K-12 teachers, as well as more likely to be recognized and rewarded for such efforts.

Another possibility is that the funding agencies would be best served by targeting a different type of STEM faculty, such as those who are not on a tenure track, because the pressure for tenure and promotion poses barriers to involvement in K-20 reforms. This is especially relevant when considered in light of the realization that the fast expansion of non-tenure-track faculty and part-time faculty is causing the landscape of IHE faculty to change rapidly (Schuster & Finkelstein, 2006). According to the Integrated Postsecondary Education Data System (IPEDS) Fall Staff Survey for 2003, 34.8 percent of all full-time faculty are off-track faculty, and they represent 58.6 percent of all newly hired full-time faculty. In addition, part-time faculty already account for half of the academic workforce. It is important to note the marked differences within the system. Research and doctoral institutions and the more selective liberal arts colleges, while increasingly resorting to contingent staff, still retain a large majority of full-time tenured or tenure-track faculty. Nevertheless, without the pressure of obtaining tenure through research and publications, non-tenure-track faculty members may be a resource that could be more fully utilized in the future.

References

Abbott, C., Daft, J., Flanagan, P., Standing, K., Tenenbaum, E., von Glatz, A., & Wiles, R. (1992). *Reaching for college*. Vol. 1/2, *Directory of college-school partnerships*. Washington, DC: Office of Policy and Planning, U.S. Department of Education.

Alshare, K.A., Wenger, J., & Miller, D. (2007). The role of teaching, scholarly activities, and service on tenure, promotion and merit pay decisions: Dean's perspective. *Academy of Educational Leadership Journal, 11,* 1, 53-68.

Boyer, E.L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Chait, R. (1998). *Ideas in incubations: Three possible modifications to traditional tenure policies*. American Association for Higher Education, New Pathways Working Paper Series, Inquiry No. 9. Washington DC: American Association for Higher Education.

Colbeck, C. (1994, November). *The contexts of academic work: What matters to faculty* (ED 375732). Paper presented at the 19th annual meeting of the Association for the Study of Higher Education, Tucson, AZ.

Diamond, R. (1999). *Aligning faculty rewards with institutional mission: Statements, policies and guidelines*. Bolton, MA: Anker Publishing Company, Inc.

Gray, P., Diamond, R., & Adam, B. (1996). *A national study on the relative importance of research and undergraduate teaching at colleges and universities*. Syracuse, NY: Syracuse University.

Jaeger, A.J., & Thornton, C.H. (2008). Neither honor nor compensation: Faculty and public service. *Educational Policy, 20,* 2, 345-366.

McDowell, G.R. (2001). *Land-grant universities and extension into the 21st century: Renegotiating or abandoning a social contract*. Ames, IA: Iowa State University Press.

Miles, M.B., & Huberman, A.M. (1994). *An expanded sourcebook: Qualitative data analysis*. Thousand Oaks, CA: Sage Publications.

O'Meara, K. (2005). Reconsidering the effects of encouraging multiple forms of scholarship: Findings from a national survey of chief academic officers. In K.A. O'Meara & R.E. Rice (Eds.), *Encouraging multiple forms of scholarship: Experienced voices*. Washington, DC: American Association for Higher Education.

Patton, M.Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.

Schuster, J., & Finkelstein, M. (2006). On the brink: Assessing the status of the American faculty. *Thought & Action*, Fall, 51-62.

Senge, P.S. (1990). *The fifth discipline*. New York: Currency Doubleday.

Sirotnik, K., & Goodlad, J. (1998). *School university partnerships in action*. New York: Teachers College Press.

Teitel, L. (1999). Looking toward the future by understanding the past: The historical context of professional development schools. *Peabody Journal of Education, 74,* 3&4, 6-20.

Xiaodong Zhang is senior study director, Education Studies, Westat, 1600 Research Boulevard, Rochville, MD 20850-3195. Correspondence concerning this article can be sent to <xiaodongzhang@westat.com> or by phone at (240) 453-2671.

Joseph McInerney is senior research, Education Studies, Westat, 1600 Research Bouldvard, Rochville, MD 20850-3195.

Joy Frechtling is vice president, Westat, 1600 Research Boulevard, Rockville, MD 20850-3195.

This research is supported by funding from the Research, Evaluation and Technical Assistance grant under the National Science Foundation's Math and Science Partnership program. We wish to acknowledge four other Westat researchers who worked on this project: Joan Michie, John Wells, Atsushi Miyaoka and Glenn Nyre, as well as 10 STEM site visitors, and 5 advisory panel members. The views expressed in this article are solely those of the authors and do not necessarily reflect those of the funders.

Angela M. Kelly, Keith Sheppard

The Relationship Between the Urban Small Schools Movement and Access to Physics Education

Although the positive and negative aspects of small schools have been well documented in some respects, this study examines one issue that has been missing from the debate, namely, the relationship between school size and physics access.

Introduction

New York City is presently undergoing a fundamental shift in secondary school organization that may have long-term ramifications for academic proficiency and college preparedness. One of the key indicators of this relationship may be access to high school physics. The relatively new *Small Schools Initiative*, a centerpiece of the New York City Department of Education's (NYCDoE) school reform efforts, began as a replacement strategy for large, failing high schools. This policy reform intended to address poor educational outcomes in historically underserved communities. Since 2002, approximately 290 new small secondary schools have opened, either alone in buildings or sharing buildings with other public schools. The Chancellor of Schools has indicated that he hopes to have 25% of the city's 350,000 high school students enrolled in small schools in the future (Greene & Symonds, 2006).

The movement has met with some success. The NYCDoE recently announced that the graduation rate in 2007 for the new small schools had exceeded 70% for the second consecutive year, compared with the citywide graduation rate of 59% (NYCDoE, 2008a). Notably, more than two-thirds of students in these graduating classes entered their new schools performing below grade-level, and more than 90% were underrepresented minority students. Many of these schools have replaced those where the collective graduation rate in 2002 was 35% (NYCDoE, 2008b). The movement has received external funding from several philanthropic groups, such as the Gates Foundation (Bill & Melinda Gates Foundation, 2005) and the Carnegie Corporation of New York.

While much of the research on small schools has been positive, focusing on safer physical and emotional learning environments, the relationship of the movement to science education has rarely been considered. Smaller schools tend to constrict curricular offerings, which can influence student access to more advanced courses (Weingarten, 2004). The purpose of this study was to identify potential correlations between the small schools movement and student access to and enrollment in secondary physics.

Secondary school restructuring has had documented positive and negative effects.

The availability of physics courses is important for several reasons. Physics is an essential component of scientific literacy (National Research Council, 1995), and it also acts a gateway course for post-secondary study in science, medicine, and engineering (Madigan, 1997; Tyson, Lee, Borman & Hanson, 2007). Limited physics access typically results in diminished scientific proficiency and college science readiness (American College Testing [ACT], 2006; National Assessment of Educational Progress, 2005). Students who have taken physics are much more likely to attend four-year colleges

(Hoffer, 1995). Indeed, completion of a high school physics course has been shown to increase retention in postsecondary STEM study, particularly for underrepresented minorities (Tyson et al., 2007).

Research Questions

This study sought to examine physics offerings and enrollments, as well as teacher certification data, in small schools (<600 students) throughout New York City during the 2004-2005 academic year; results were compared to those for mid-sized (600-1200 students) and large (>1200 students) high schools. The following overarching research question was considered: What is the relationship between school size and access to physics study, an established indicator of college preparedness?

In addition, the study led us to examine several sub-questions: (a) How does physics course availability relate to school size when compared to other variables?; (b) Among schools that offer physics, which types of physics courses are available?; and (c) What is the proportion of physics-certified teachers, and how does this relate to school size?

These research questions were explored utilizing data from the NYCDoE, as well as survey responses obtained directly from school administrators. The analysis first considers current research on small schools, the status of physics availability in New York City's schools, and recent patterns of physics teacher quality. Next, the quantitative methodology and research design are described. Finally, results from the survey and a discussion of the potential relationship between the small schools initiative and physics education are presented.

Review Of Literature

Prior studies provide the framework for the analysis of the research questions. Secondary school restructuring has had documented positive and negative effects. Positive effects are mainly related to learning environment, student engagement, and graduation rates. The negative effects include curricular limitations, although physics has not been specifically addressed in existing research. Finally, the availability of physics has been typically lacking in urban schools, which is further compounded by a shortage of physics-certified teachers.

Arguments that Support the Establishment of Small Secondary Schools

Recently, New York City and many other urban areas have undergone a shift to smaller secondary schools. Larger high schools have often been viewed as detached, segregated institutions, which unwittingly provide unequal learning opportunities (Ilg & Massucci, 2003; Lipman, 1998). Administrative and instructional practices in such schools have tended to isolate students, who do not have much connection with teachers and staff (Darling-Hammond, Ancess & Ort, 2002).

A systemic shift to smaller high schools in New York City came with the hope of greater individual success. One obvious advantage is the opportunity for students to participate in a close community with shared goals (Lipman, 1998; Osterman, 2000). Students cannot remain anonymous and exhibit fewer behavioral problems and less frequent absenteeism (Darling-Hammond et al., 2002; Stiefel, Berne, Iatarola & Fruchter, 2000). Student advisement has been more meaningful and consistent (Shear et al., 2008), and more students have attended college (Schneider, Swanson, & Riegle-Crumb, 1998).

A systemic shift to smaller high schools in New York City came with the hope of greater individual success.

Most importantly, small schools appear to have promoted effectiveness and equity, aiming to decrease the gap in opportunities for students of different socioeconomic status (Lee, Smith, & Croninger, 1996), thereby lessening its effect on student performance (Copland & Boatright, 2004; Darling-Hammond et al., 2002). In traditional high schools, students are frequently separated according to ability, which results in social stratification (Ready & Lee, 2008).

Small schools have been politically popular in New York City, mainly because most realize that large high schools are not meeting the needs of the city's students. However, this initiative is rather new—approximately 290 schools since 2002 (NYCDoE, 2008b)—so the political climate may change if these schools do not produce measurable results that prove their effectiveness (Bloomfield, 2005).

Potential Limitations of Small Schools

Despite the promise of drastic educational improvement through a shift to smaller schools, there have been several issues that suggest this promise may not be kept. These schools typically offered curricular uniformity as a means for greater

achievement (Monk & Haller, 1993). The rationale is that students adhere to a more academic course of study if there are fewer electives. Although this has resulted in increased graduation rates (Copland & Boatright, 2004; Greene & Symonds, 2006), students in smaller schools have not had as many opportunities to participate in advanced courses (Miner, 2005). Smaller schools with reduced staffs have fewer course options (Monk & Haller, 1993; Ravitch, 2005a), which are usually targeted towards mid-level students (Lee et al., 2000).

Small schools have been politically popular in New York City, mainly because most realize that large high schools are not meeting the needs of the city's students.

Although Lee et al. (1996) elaborated on the harmful nature of course differentiation between schools, particularly for disadvantaged students, they focused on achievement rather than access. Gamoran (1996) pointed out that substantial between-school variation in opportunity must indirectly affect achievement. Few studies have substantiated the limitations of the small school structure (Herszenhorn, 2005; Ravitch, 2005b). However, an analysis of small schools data in NYC showed that small school students performed worse academically than those in larger high schools (Iatarola, Schwartz, Stiefel & Chellman, 2008). Bill Gates recently expressed his disappointment in the results of his small schools investment, stating that many did not improve achievement (Gates, 2009).

Another potential unintended consequence of small schools has been a lack of teacher support in implementing curricular reforms. Darling-Hammond (1997) suggested that these schools better served students by employing fewer administrators, since teachers take on many of their duties (counseling, discipline, etc.). However, this assumes that teachers have the expertise required to fulfill these roles. A traditional content supervisor, usually nonexistent in small schools, is instrumental in helping teachers develop sound pedagogical practices (Miner, 2005; Wyse, Keesler, & Schneider, 2008).

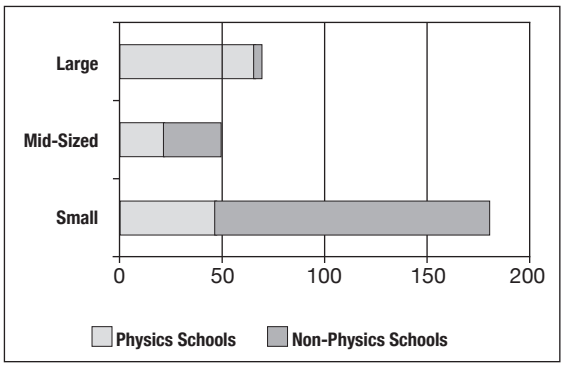
Although the positive and negative aspects of small schools have been well documented in some respects, this study examines one issue that has been missing from the debate, namely, the relationship between school size and physics access. Do small schools, which have been successful in improving graduation rates and creating a sense of community, prioritize advanced science courses? It is first necessary to examine the current status of physics in U.S. schools, and how physics access has been a particular concern for students in urban settings.

Contemporary Availability of Physics in Urban Secondary Schools

In a previously published report, the authors noted that approximately 21% of NYC high school graduates have taken physics (Kelly & Sheppard, 2008), compared with the national average of 33% for public schools (Neuschatz, McFarling, & White, 2008). The city also lagged behind the New York State's physics enrollment of 31% (New York State Education Department [NYSED], 2004). An analysis of the distribution of physics in city schools showed that access to physics was not equitably distributed—55% (164 of 298) of the surveyed New York City high schools did not offer it. This stands in stark contrast to the figures reported nationally, where physics was offered in 89% of high schools (Neuschatz et al., 2008). Additionally, race, socioeconomic status, and prior academic achievement appeared to be related to physics availability (Kelly & Sheppard, 2009).

The size of city high schools was an important factor in predicting whether or not physics was offered (see Figure 1). The vast majority (96%) of large high schools (n> 1200 students) offered physics while 45% of mid-sized schools (600<n<1200) and only 26% of the small schools (n<600) did (Kelly & Sheppard, 2008). The present study examines this trend in more detail in order to evaluate factors related to physics availability in small schools.

Figure 1: Physics availability in terms of school size in New York City, 2004-05 (n=298).



Physics Teacher Quality and Certification

The availability of qualified physics instructors has been a barrier to

Physics courses are not available in a large proportion of high schools in NYC, and this negatively affects approximately 75,000 students.

ensuring equitable opportunities in physics in urban districts (Brumberg, 2000; Hodapp, Hehn, & Hein, 2009; Ingersoll, 1999). The Council of Chief State School Officers [CCSSO] noted that teacher quality is crucial in preparing scientifically literate graduates (2007). The National Science Foundation (2008) reported that 78% of all secondary physical science teachers were certified in their respective disciplines, while CCSSO reported in 2007 that science teacher certification rates in physics had declined slightly since 1996. This may have been influenced by the rising number of secondary physics students nationwide, which reached a high of 1.1 million in 2005 (Neuschatz et al., 2008). Rapidly increasing enrollments have resulted in some states hiring less-qualified teachers to meet demand.

The academic background required to teach physics in New York State has shifted in recent years. Presently, candidates must have a bachelor's degree or 30 credits in physics, and must also pass several standardized tests (NYSED, 2009b). Prior to 2004, the minimum content requirement for physics teachers was 18 credits in physics. Consequently, all New York physics teachers should have at least a minor in physics.

The situation is, however, complicated by the practice in New York of "incidental teaching," through which school districts can allow teachers to teach one course out of their certification area. The number of teachers teaching physics through this method is unknown, though New York City has fewer qualified teachers than the rest of state (Brumberg, 2000). During the 2007-08 academic year, 16% of NYC science teachers did not have appropriate disciplinary certification; teachers who were not "highly qualified" staffed 4.8% of all biology classes, 6.0% of chemistry classes, 16.5% of Earth science classes, and 10.9% of physics classes (NYSED, 2008). According to a variety of reports the recruitment and retention of qualified physics instructors is a necessity to maintain viable secondary physics programs (The City Council of New York, 2004; Monk, 1994; Osbourne, 2003).

Methodology

As part of a larger study on physics education in New York City during the 2004-2005 academic year, information on physics enrollment and course offerings was collected using a written survey. The survey was necessary because the data could not be obtained from the NYCDoE; there was no information on physics availability and schools had to be contacted individually. Ultimately, data were collected from 298 out of the 316 schools (94%).

The information collected from the surveys included the number of physics students, the number of sections offered, the certification status of physics teachers, and the type(s) of physics courses available. Three major types of physics courses were identified:

1. *Regents Physics* is a traditional college-preparatory physics course based on the standardized New York State Physical Setting Curriculum (NYSED, 2009a). Regents Physics is normally taken by students after they have completed Regents courses in Biology (known as Living Environment), Earth Science, and Chemistry. Students are required to complete at least 1200 minutes of laboratory work before they can take the examination. Students must pass at least 5 Regents examinations in order to graduate from high school, one of which must be a science.
2. *Advanced Placement Physics* is a college-level physics course based on a College Board curriculum (College Board, 2009). This course can be either algebra-based (AP Physics B) or calculus-based (AP Physics C). The curriculum is standardized, and schools that offer the course must also complete a curricular audit to ensure compliance. A grade of 3 or better on the culminating AP Physics Exam indicates passing proficiency.
3. *Non-Regents Physics* is a thematic-based, less mathematically-oriented physics course that is often taught with the text *Active Physics* (Eisenkraft, 1998) in city schools. This course is more conceptual in nature with less challenging mathematical applications.

The survey specifically asked if physics teachers were certified in physics, as opposed to other scientific disciplines.

The dependent variable in this study was whether or not a school offered physics. Several school-level characteristics were the independent variables in the study, and data for these variables were obtained from

the Annual School Reports Cards (NYCDoE, 2006). The dependent variables included: school enrollment, graduation rate, percentage of students attending two- and four-year colleges, average SAT Math scores (a measure of mathematical proficiency), and passing rates on the biology and chemistry state standardized exams. The variables are defined in Table I. Not all of the data were available for each of the schools since some Annual School Reports were incomplete.

To answer the research questions, schools were grouped into three categories by population: large schools (>1200 students), mid-sized schools (600-1200 students), and small schools (<600 students). Similarly defined groups were used by Crocco and Thornton (2002) in their descriptive study of social studies in NYC schools, although a mid-sized range was created in this study to bridge the difference between large and small schools. Since 70% of U.S. high school students were enrolled in schools with more than 1000 students, and 50% are in schools with more than 1500 students (U.S. Department of Education, 2008), the researchers designated a mid-sized category to see if three categories could provide clearer insights. The average small school in NYC enrolled 286 students, the average mid-sized school had 816 students, and the average large school had 2892. The researchers felt that these three categories represented very different types of learning environments. Statistical analyses revealed patterns in physics course offerings, teacher certification, and school size.

Correlations between physics access and several organizational variables were examined. The differences between the variable means for physics and non-physics schools were calculated using independent-samples t-tests, based on a confidence level of 95%. Effect size was measured using Cohen's d as a benchmark for large (0.8), medium (0.5), and small (0.2) effect sizes (Cohen, 1988). A multivariate analysis of variance (MANOVA) with Bonferroni confidence intervals was performed to determine whether these predictor variables retained their significance when combined in a multivariate model (p <.001). Finally, a binary logistic regression was performed with the independent variables in order to determine how much they accounted for the variance between physics and non-physics schools. The R-squared values were computed for various combinations of the variables to identify the highest correlation with physics availability.

Table 1: Independent variables related to physics availability.

Independent Variables	How Defined
School enrollment	The total school population in grades 9-12, including special education students, both mainstreamed and self-contained. All other grades within the school were eliminated from the total.
Graduation rate	This rate is the percentage of students who began as freshman and graduated from the same school. The graduation rate for the Class of 2004 was based on students who matriculated in the school in 2000. The number included students who earned a GED or a diploma with or without a Regents endorsement.
Percentage of students attending two- and four-year colleges	This variable was a composite of the percentages of students who attended either 2-year or 4-year colleges upon graduation.
Average SAT Math score	The SAT is a standardized reasoning test developed by the College Board that has 3 sections: mathematics, verbal, and writing. This test is required by many colleges for admission.
Passing rates on New York State standardized exams in Biology and Chemistry	Science achievement was a broad indicator of the level of science performance within each school. This variable was reported by two measures: the percentage of students passing the Living Environment (Biology) Regents and Chemistry Regents examinations. The passing score was 65%.

Limitations

Several limitations relating to the research design must be acknowledged. First and foremost, physics availability and teacher certification data were self-reported, with no means of triangulating the results. Since schools have an interest in hiring highly qualified teachers in all content areas (No Child Left Behind Act, 2002), certification status figures may be inflated. By state regulation, secondary teachers in New York State must teach 80% of their classes within their area of certification, with no more than five classroom hours per week teaching in other disciplines (NYSED, 2009b), which is difficult when many of the surveyed city schools offered only one section of physics. In addition, administrators and teachers who reported enrollments may have estimated numbers of students,

resulting in some inaccuracies in the tabulations. Finally, some of the Annual School Reports did not report complete data for each school. However, this did not affect significance, and it was factored into effect size calculations.

Results

The different sized schools showed clear trends in physics enrollment distribution by course type (see Figure 2 and Table 2). Small schools had higher numbers of students enrolled in Non-Regents Physics (53%) compared to Regents Physics (47%); notably, not a single small school was able to offer Advanced Placement (AP) Physics. The great majority of physics courses offered at mid-sized schools were Regents-based (87% of all physics students), with 8% of their students taking Non-Regents and 5% taking AP courses. The types of physics courses offered at large schools had a similar distribution, with 83% of students taking Regents Physics, and more students enrolled in Non-Regents or AP (9% and 8%, respectively) than in mid-sized schools. The large school trend reflected the national trend described by the American Institute of Physics, where increased numbers of students were taking Non-Regents/Conceptual Physics and AP Physics (Neuschatz et al., 2008). The percentage of students in traditional first-year physics courses has been declining as a result.

Figure 2: Distribution of physics enrollment by physics type and school size, 2004-05.

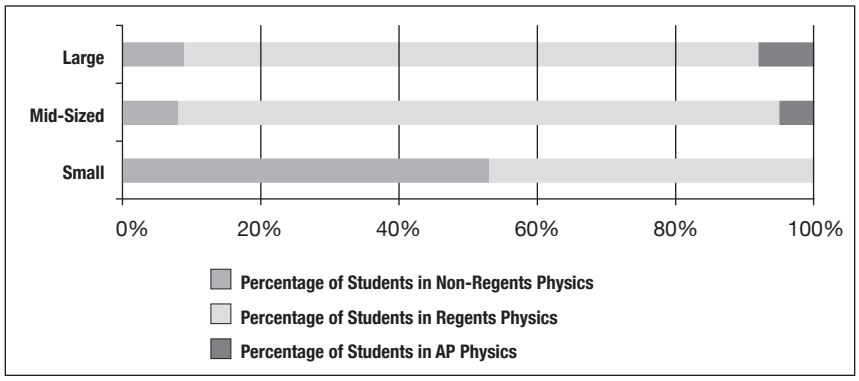


Table 2 also displays the percentage of physics students in each category of school size. Even though the majority of small schools did not offer physics at the time of the study, small schools still had a higher percentage of students taking physics (5.7%) than mid-sized (5.6%) and large schools (5.0%).

The Relative Importance of School Size in Predicting Physics Access

Independent samples t-tests were conducted to assess differences between physics and non-physics schools in the following variables (school characteristics):

- School size/enrollment
- Average SAT Math score
- Graduation rate
- Passing rates on chemistry standardized test
- Passing rates on biology standardized tests
- Percentage of college bound students

Table 3 shows that physics and non-physics schools differed significantly on all variables when considering t-values and effect sizes. Non-physics schools tended to be quite small (372 students), while physics school tended to be large (1579 students). Physics schools had higher SAT Math scores, higher graduation rates, higher passing rates on both chemistry and biology standardized tests, and higher college attendance rates. All of the tested variables had either large effect size (enrollment, SAT Math, graduation rate) or medium

Table 2: Numerical distribution of physics enrollment by physics type and school size, 2004-05.

Variable	Small Schools	Mid-Sized Schools	Large Schools
Total number of students enrolled	51,162	41,493	194,207
Number of students taking any physics course	2,937	2,303	9,695
(% of total enrollment)	(5.7%)	(5.6%)	(5.0%)
Number of students in Regents Physics	1380	2,004	8046
(% of physics students)	(47%)	(87%)	(83%)
Number of students in Non-Regents Physics	1557	184	873
(% of physics students)	(53%)	(8%)	(9%)
Number of students in AP Physics	0	115	776
(% of physics students)	(0%)	(5%)	(8%)

Table 3: Comparison of effect sizes of characteristics for physics and non-physics schools. (p < .01)

Variable	t	Degrees of freedom	Effect size	Mean of physics schools	Meanof non-physics schools	Mean difference	Standard error difference
Enrollment	11.3	296	1.32 (large)	1579	372	1207	112.3
SAT Math	5.9	160	0.93 (large)	449	389	60	10.3
Graduation rate	5.7	175	0.86 (large)	64%	47%	17%	3.0
Percentage passing chemistry standardized tests	4.4	140	0.74 (medium)	49%	27%	22%	4.9
Percentage passing biology standardized tests	4.2	198	0.60 (medium)	57%	40%	17%	3.9
Percentage of students college-bound	3.6	164	0.56 (medium)	74%	59%	15%	4.1

effect size (percentage passing Living Environment and Chemistry Exams, percentage attending college). According to effect size, enrollment had the greatest correlation with physics availability.

The variables that describe academic performance were also examined by school size to see if there were a difference in the means. As reported in Table 4, students in small schools had the lowest mean SAT Math scores (410) when compared to mid-sized (419) and large (454) urban high schools. Their passing rates on Chemistry (35%) and Living Environment (46%) Regents Exams were lower than those for large schools (44% and 53%, respectively), but slightly higher than mid-sized schools (30% and 45%). The graduation (58%) and college attendance (67%) rates for students in small and mid-sized schools (57% and 66%, respectively) were approximately the same, but the graduation rate for large schools was slightly higher (60%) and the percentage of college-bound students was slightly lower for large schools (64%).

Once the significance and effect size of the individual variables were determined, a multivariate analysis of variance (MANOVA) was performed using Bonferroni confidence intervals. The joint multivariate Bonferroni approach sets up 95% confidence intervals around the parameter coefficients. All of the previously tested individual variables remained significant (Wilk's1 = .651, $p < .001$) in the multivariate model, as indicated by the confidence intervals in Table 5.

Finally, the six numerical independent variables were analyzed using binary logistic regression, in order to determine to what extent they accounted for the variance between physics and non-physics schools. When combined into one equation, the six variables had an R-squared value of 0.614 (p <.01), indicating that they accounted for 61.4% of the variance between the two types of schools. Through backward stepwise regression, it was apparent that

Table 4: Comparison of means for enrollment and academic performance among small, mid-sized, and large high schools.

Variable	Small Schools	Mid-Sized Schools	Large Schools
Enrollment	286	816	2892
SAT Math	410	419	454
Graduation rate	58%	57%	60%
Percentage passing chemistry standardized tests	35%	30%	44%
Percentage passing biology standardized tests	46%	45%	53%
Percentage of students college-bound	67%	66%	64%

Table 5: Analysis of variance of independent variables using Bonferroni confidence intervals.

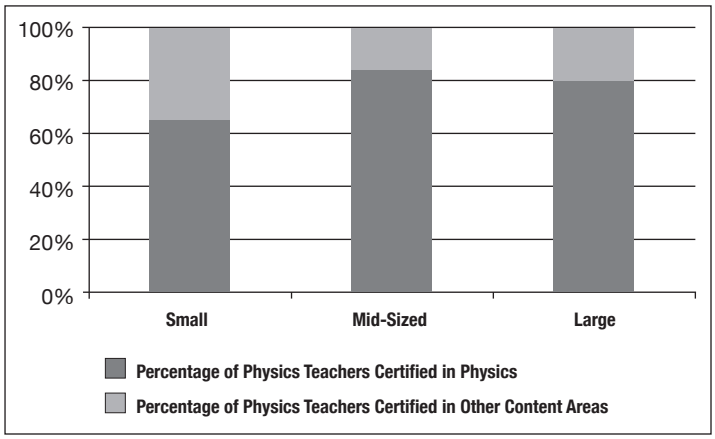
Variable	t	Standard error	95% confidence interval of the difference	
			lower	upper
Enrollment	5.0	280	639	2177
SAT Math	4.2	14.6	21.7	102
Graduation rate	3.7	4.23	4.12	27.3
Percentage passing chemistry standardized tests	3.9	5.92	6.64	30.0
Percentage passing biology standardized tests	3.8	5.49	5.94	36.0
Percentage of students college-bound	3.1	5.48	2.16	32.2

biology and chemistry standardized test scores and college attendance rate did not contribute significantly to the model, as expected when considering effect size. After eliminating these variables, enrollment, SAT Math scores, and graduation rate accounted for 51.1% of the variance (R-squared = 0.511). The model was further analyzed by examining various iterations of independent variables to predict physics availability, and each time school enrollment was the most significant predictor, followed by SAT Math scores and graduation rate. For example, enrollment and graduation rate predicted 49.7% of the variance, enrollment and SAT Math scores predicted 48.4%, and graduation rate and SAT Math scores predicted 31.5%. This again indicates that school size is the single largest correlating factor when considering whether or not physics was offered.

Physics-Certified Teachers in Small Schools

The distribution of physics-certified teachers was also examined in the context of school size (Figure 3). A total of 114 physics teachers in 91 schools responded to the survey question regarding certification. Large and mid-sized schools had an advantage over small schools in terms of the percentage of physics teachers who were certified in the field: 80% of physics teachers in large schools and 84% of teachers in mid-sized schools were certified in physics, while just 65% of physics teachers in small schools were certified in the subject. The sample size is smaller than that of the overall study because administrators often did not report the certification status of their physics teachers in the survey response.

Figure 3: Location of physics-certified teachers by school size (n=91 schools), 2004-05.



Discussion

A basic understanding of physics is critically important for comprehending contemporary scientific and technological issues. It is problematic that 79% of city students graduate high school without having taken a physics course. Physics courses are not available in a large proportion of high schools in NYC, and this negatively affects approximately 75,000 students. The availability and organization of the sciences in small schools needs to be reconsidered if future generations of students are not to be disenfranchised by restricted physics access.

That so few small schools offer physics (only 26%) is a troubling trend, because the city plans on increasing the number of small schools, particularly in high poverty areas such as the Bronx. Although small schools have improved school climate and student retention, the question of access to advanced science needs to be answered. Limited physics opportunities are detrimental for students who wish to pursue post-secondary science (ACT, 2006; Tyson et al., 2007). If New York City continues to open more small schools, the number of students who have graduated after taking physics may actually decrease.

A positive note regarding physics in the small schools is that a higher percentage of the overall small school population (5.7%) was enrolled in physics than mid-sized (5.6%) and large schools (5.0%). This suggests that when small schools do offer physics, the curricular uniformity results in a larger proportion of students taking physics than in a typical large high

school, where many students can elect not to take it. This is an encouraging trend, yet the opportunity needs to be expanded to the other 74% of small schools that do not offer physics. If this were the case, small schools could be instrumental in boosting physics enrollment and preparing more students for advanced STEM study.

When examining academic variables related to physics availability, there are some noticeable differences between physics and non-physics schools, and among schools of different sizes. Physics schools typically reported higher SAT Math scores, passing rates on standardized science tests, and higher graduation and college attendance rates. Small schools reported lower SAT Math than mid-sized and large schools, and considerably lower science scores than large high schools. This suggests that prior science and math achievement may negatively impact whether or not schools choose to offer physics. Although small schools have the highest college attendance rates of the three school types, their students seem to perform at a lower level than those in larger schools. Further research is needed to understand how small schools might leverage improved student retention to boost achievement.

School size also appears to be a factor in the types of physics courses offered. Physics courses offered in large NYC high schools mirrored the national trend (Neuschatz et al., 2008) in that some students enrolled in Conceptual and Advanced Placement Physics, while the majority enrolled in Regents Physics. However, slightly more than half of small schools offered Non-Regents Physics, and the rest offered Regents Physics. This suggests two key points for consideration. First, the small school structure seems particularly conducive to teaching Non-Regents Physics. This conceptual course, which is less mathematical in nature than a standard college-prep physics course, is well suited to a heterogeneous population with varying science backgrounds. It can be taught to younger students with relative ease. Since small schools have more course uniformity, such courses are desirable. However, the question remains whether students are adequately prepared for college physics after a single Conceptual Physics course; ideally, this course could be the first of a 2-course physics sequence. Secondly, AP Physics was only available in 20 of 298 (6.7%) schools, all with more than 800 students. A larger student population seems necessary to support study of advanced college-level physics. However, small schools could look at ways of offering higher-level physics, such as combining interested students from different schools within the same building or neighborhood, or partnering with nearby colleges.

Physics appears to be a barometer of equity when considering science course taking opportunities.

The number of physics teachers certified in physics appears to vary with school size. In small schools, only two-thirds of physics teachers held physics licenses, less than large (80%) and mid-sized (84%) schools. Science teachers in small schools often have to teach outside of their certification because there are fewer faculty. Further research is needed to determine the impact of certification status on student learning in physics classrooms.

Implications

The political leadership in New York City enthusiastically supports the expansion of the small schools initiative to improve graduation rates and other academic outcomes. However, the limited availability of physics in such schools is a curricular restraint that warrants careful consideration. The successful completion of a high school physics course provides an advantage for students who plan to attend college, particularly those who wish to study STEM-related disciplines Students are more likely to persist if they have taken a physics course. The causes for this relationship are beyond the scope of this study. However, physics appears to be a barometer of equity when considering science course taking opportunities. It is a source of educational capital, in that it provides authenticity and status, and it is often a requirement for admission to competitive colleges. Since smaller schools seem to have considerable difficulty in offering physics, further expansion should be contingent upon realistic policy proposals that aim to equalize access to it.

The question of physics availability is part of the larger issue of transparency. Although there have been some studies related to the academic performance of small school graduates, few have examined students' achievement in terms of opportunity-to-learn considerations. There are fewer advanced courses, but what does this mean in terms of student outcomes? The investment of capital in the small schools initiative necessitates accountability structures that ensure students are receiving more than just

an adequate education (Rebell & Wolff, 2008). Reliable data on course availability should identify potential curricular shortfalls and propose solutions for improvement. Smallness alone does not guarantee excellence; schools must hold themselves accountable to rigorous academic standards (Copland & Boatright, 2004; Iartarola et al., 2008). These standards should be communicated to parents and students so that opportunities are clearly understood before students enroll.

Examining course offerings and teacher certification are two ways to analyze how well these schools are meeting the needs of their students. Small schools are retaining students, a positive outcome, but their effectiveness in terms of availability of advanced courses and highly qualified teachers (and, by extension, achievement) needs further examination. As with most academic innovations, educators must find ways to maintain the advantages while minimizing harmful consequences. With the development of creative solutions to improve access to, and success in, advanced science, small schools have tremendous potential to enhance educational opportunities for urban children.

References

American College Testing. (2006). Developing the STEM education pipeline. Washington, DC: American College Testing (ACT), retrieved 1/5/2009, from <www.act.org/path/policy/index.html>.

Bill & Melinda Gates Foundation. (2005). Multiple Pathways to Graduation Initiative, retrieved 6/1/2007, from <http://www.gatesfoundation.org/education/transforminghighschools>.

Bloomfield, D.C. (2005). High school reform: The downside of scaling-up. Politics of Education Association Bulletin, 30(1), 6-9.

Brumberg, S.F. (2000). The teacher crisis and educational standards. In D. Ravitch & J.P. Viteritti (Ed.), City schools: Lessons from New York (pp.141-165). Baltimore: Johns Hopkins University Press.

The City Council of New York. (2004). Lost in space: Science education in New York City public schools. New York: The City Council of the City of New York Committee on Education.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. New York: Lawrence Erlbaum.

College Board. (2009). College Board AP Physics course description, retrieved 3/1/2009, from <www.collegeboard.com/prod_downloads/ap/students/physics/ap-cd-physics-0708.pdf>.

Copland, M., &Boatright, E. (2004). Leading small: Eight lessons for leaders in transforming comprehensive high schools. Phi Delta Kappan, 85(10), 762-770.

The Council of Chief State School Officers. (2007). State indicators of science and math education. Washington, DC: The Council of Chief State School Officers, retrieved 3/6/2009, from <http://www.ccsso.org/projects/Science_and_Mathematics_Education_Indicators/>.

Crocco, M. S., & Thornton, S.W. (2002). Social studies in the New York City public schools: A descriptive study. Journal of Curriculum and Supervision, 17(3), 206-231.

Darling-Hammond, L. (1997). The right to learn: A blueprint for creating schools that work. San Francisco: Jossey-Bass.

Darling-Hammond, L., Ancess, J., & Ort, S.W. (2002). Reinventing high school: Outcomes of the coalition campus school project. American Educational Research Journal, 39(3), 639-673.

Eisenkraft, A. (1998). Active physics. New York: Herff-Jones.

Gates, W. (2009). 2009 Annual letter from Bill Gates: U.S. education, retrieved 2/28/2009, from <http://www.gatesfoundation.org/annual-letter/Pages/2009>.

Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. Educational Evaluation and Policy Analysis, 18(1), 1-18.

Greene, J., & Symonds, W.C. (2006, June 26). Bill Gates gets schooled: Why he and other execs have struggled in their school reform efforts, and why they keep trying. Business Week, pp.64-70.

Herszenhorn, D. M. (2005, November 8). In New York's inaugural smaller schools, "a good year and a tough year". The New York Times.

Hodapp, T., Hehn, J., & Hein, W. (2009). Preparing high-school physics teachers. Physics Today, 62(2), 40-45.

Hoffer, T. (1995). High school curriculum differentiation and postsecondary outcomes. In P.W. Cookson & B. Schneider (Eds.), Transforming schools (pp.371-402). New York: Garland.

Iatarola, P., Schwartz, A.E., Stiefel, L., & Chellman, C.C. (2008). Small schools, large districts: Small-school reform and New York City's students. Teachers College Record, 110(9), 1837-1878.

Ilg, T. J., &Massucci, J.D. (2003). Comprehensive urban high school: Are there better options for poor and minority children? Education and Urban Society, 36(1), 63-78.

Ingersoll, R. (1999). The problem of underqualified teachers in American secondary schools. Educational Researcher, 28(2), 26-37.

Kelly, A.M., & Sheppard, K. (2008). Newton in the Big Apple: Access to high school physics in New York City. *The Physics Teacher*, *46*(5), 280-283.

Kelly, A.M., & Sheppard, K. (2009). Secondary physics availability in a U.S. urban district: Issues related to academic achievement and course offerings. *American Journal of Physics, 77*(10), 902-906.

Lee, V. E., Smerdon, B.A., Alfred-Liro, C., & Brown, S.L. (2000). Inside large and small high schools: Curriculum and social relations. Educational Evaluation and Policy Analysis, 22(2), 147-171.

Lee, V. E., Smith, J.B., & Croninger, R.G. (1996). Understanding high school restructuring effects on the equitable distribution of learning in mathematics and science. Madison, WI: Center on Organization and Restructuring of Schools.

Lipman, P. (1998). Race, class, and power in school restructuring. Albany: State University of New York Press.

Madigan, T. (1997). Science proficiency and course-taking in high school: The relationship of science course-taking patterns to increases in science proficiency between 8th and 12th grades. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Educational Statistics.

Miner, B. (2005). The Gates Foundation and small schools. Rethinking Schools Online, 19(4), retrieved 10/25/2008 from <http://www.rethinkingschools.org/archive/19_04/gate194.shtml>.

Monk, D.H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. Economics of Education Review, 13, 125-145.

Monk, D.H., & Haller, E.J. (1993). Predictors of academic course offerings: The role of school size. American Educational Research Journal, 30, 3-21.

National Assessment of Education Progress. (2005). America's report card: Results from the 2005 transcript study, retrieved 11/12/2008 from <http://nationsreportcard.gov/hsts_2005/>.

National Research Council. (1995). National science education standards. Washington, DC: National Academy Press.

National Science Foundation (2008). Science and engineering indicators, 2008, retrieved 2/28/2009 from <http://www.nsf.gov/statistics/seind08/pdf/c01.pdf>.

Neuschatz, M., McFarling, M., & White, S. (2008). Reaching the critical mass: The twenty year surge in high school physics. College Park, MD: American Institute of Physics.

New York City Department of Education. (2006). Annual school reports, 2004-05, retrieved 4/25/2006 from <http://schools.nyc.gov/daa/schoolreports/>.

New York City Department of Education. (2008a). Annual school reports, 2006-07, retrieved 10/2/2008 from <http://schools.nyc.gov/daa/schoolreports/>.

New York City Department of Education. (2008b). Chancellor Klein announces opening of 52 new schools for September, retrieved 3/2/2009 from <http://schools.nyc.gov/Offices/mediarelations/NewsandSpeeches/2007-2008/nuschools.htm>.

New York State Education Department. (2004). New York: The state of learning, A report to the Governor and the Legislature on the educational status of the state's schools, retrieved 3/1/2009 from <http://www.emsc.nysed.gov/irts/655report/2003/tableofcontents-july-2003.html>.

New York State Education Department. (2008). *Progress report on teacher supply and demand in New York State*. Retrieved August 1, 2009, from <http://www.highered.nysed.gov/tsd2008.htm>.

New York State Education Department. (2009a). Physical setting/Physics core curriculum, retrieved 3/1/2009, from <www.emsc.nysed.gov/ciai/mst/pub/phycoresci.pdf>.

New York State Education Department. (2009b). Requirements relating to teaching practice and specialized credentials, Section 80-5.3: Incidental teaching, retrieved 3/9/2009 from <http://www.highered.nysed.gov/tcert/part80-5.htm#5.3>.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).

Osbourne, J. (2003). Attitudes towards science: A review of the literature and its implications. International Journal of Science Education, 25(9), 1049-1079.

Osterman, K. F. (2000). Students' need for belonging in the school community. Review of Educational Research, 70(3), 323-367.

Ravitch, D. (2005a, March 15). Failing the wrong grades: The right—and wrong—way to improve America's public high schools. The New York Times.

Ravitch, D. (2005b, November 6). Downsize high schools? Not too far. The Washington Post.

Ready, D.D., & Lee, V.E. (2008). Choice, equity, and the schools-within-schools reform. Teachers College Record, 110(9), 1930-1958.

Rebell, M., & Wolff, J. (2008). Moving every child ahead: From NCLB hype to meaningful educational opportunity. New York: Teachers College Press.

Schneider, B., Swanson, C., & Riegle-Crumb, C. (1998). Opportunities for learning: Course sequences and positional advantages. Social Psychology of Education, 2, 25-53.

Shear, L., Means, B., Mitchell, K., House, A., Gorges, T., Joshi, A., Smerdon, B., & Shkolnik, J. (2008). Contrasting paths to small school reform: Results of a 5-year evaluation of the Bill & Melinda Gates Foundation's National High School Initiative. Teachers College Record, 110(9), 1986-2039.

Stiefel, L., Berne, R., Iatarola, P., & Fruchter, N. (2000). High school size: Effects on budgets and performance in New York City. Educational Evaluation and Policy Analysis, 22(1), 22-39.

Tyson, W., Lee, R., Borman, K.M., & Hanson, M.A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework for postsecondary degree attainment. Journal of Education for Students Placed at Risk, 12(3), 243-270.

United States Department of Education, Office of Vocational and Adult Education. (2008). School size archived information. Retrieved August 30, 2009, from <http://www.ed.gov/about/offices/list/ovae/pi/hs/schoolsize.html>.

Weingarten, R. (2004). Testimony on the establishment of new small secondary schools in New York City. The City Council of the City of New York Committee on Education, retrieved 10/20/2008, from <http://www.uft.org/news/weingarten_tes1/>.

Wyse, A.E., Keesler, V., & Schneider, B. (2008). Assessing the effects of small school size on mathematics achievement: A propensity score-matching approach. Teachers College Record, 110(9), 1879-1900.

Angela M. Kelly is assistant professor, Department of Middle and High School Education, Lehman College, City University of New York, Bronx, NY 10468-1589. Correspondence concerning this article may be sent to <angela.kelly@lehman.cuny.edu>.

Keith Sheppard is associate professor, Center for Science and Mathematics Education, Stony Brook University, Stony Brook, NY.

Daniel Z. Meyer, Leanne M. Avery

Why Inquiry is Inherently Difficult… and Some Ways to Make it Easier

The authors offer a framework that identifies two critical problems in designing inquiry-based instruction and suggest models for developing instruction that overcomes those problems.

"I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description; and perhaps I could never succeed in intelligibly doing so. But I know it when I see it …"

—Potter Stewart
Jacobellis v. Ohio

Justice Stewart's statement regarding pornography would seem to be applicable to the current state of inquiry in the science education field. We have numerous rich descriptors of inquiry in action (National Research Council, 2000; Minstrell & van Zee, 2000), as well as robust rubrics designating levels of inquiry (Herron, 1971; Wheeler, 2000; Beerer & Bodzin, 2003). In other words, we know it when we see it. But these fall short in providing teachers with the tools for how to develop inquiry-based activities. Much of the research investigating this has focused on the structural barriers (e.g. time, resources, teacher knowledge, etc.) (Anderson, 2002; Minstrell & van Zee, 2000). This research suggests areas for policy makers and teacher educators to work on, but these descriptions fall short of actually providing guidance to teachers.

While we do not aim to produce a straightforward, "cookbook" process for generating inquiry activities, we do aim to push beyond "I know it when I see it". We feel this can be done by considering the design of inquiry activities as a problem space. By exploring what makes inquiry inherently difficult, as well as three potential models that overcome these challenges, we aim to build a framework that has heuristic power. That is, it has the potential to suggest further solutions to the particular problem of designing inquiry activities. We are aiming for a middle ground between the two current extremes: something that is more general than good examples of inquiry activities, but more specific and oriented towards creating activities than outcome descriptions of inquiry in action.

Our analysis is informed both by conceptual frameworks from science studies and by the experience of facilitating a variety of science educators in developing inquiry-based instruction. This paper can be seen as a formalization of the advice we find ourselves regularly giving science educators. Our framework can be divided into two broad sections. First, we outline a problem space component—an articulation of the challenges in designing inquiry activities. Second, we provide a solution component—a series of activity types that have the potential to resolve the challenges of the problem space.

By exploring what makes inquiry inherently difficult, as well as three potential models that overcome these challenges, we aim to build a framework that has heuristic power.

Our View of Inquiry

The National Science Education Standards describe inquiry as "the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work" (NRC, 1996). Therefore, we draw heavily on studies of scientific practice to form our approach. Two concepts in particular have been useful and guide our further discussion.

The first concept is the notion that context matters. This is probably best encapsulated in Kuhn's (1970) principle of a paradigm: scientists operate in an existing framework that guides aspects of their work, such as

what counts as evidence. This effects how participants react to empirical evidence. Scientists from different fields that have points of overlap will approach common topics in different manners. For example, when results from neutrino experiments differed from current theory, different types of involved researchers questioned different parts of the theoretical framework (Pinch, 1981, 1985). But the paradigm is more than just a gauge by which to judge new work. It provides the impetuous and purposefulness that motivate researchers to take on new work in the first place. Individual pieces of scientific work (to the extent one can even define an individual piece) only have meaning in their specific context.

The second concept—interpretive flexibility—comes from sociological studies of the work of done to develop scientific and technological knowledge (Collins, 1981a, 1981b). This refers the situation in which differing conclusions can be made from the same set of empirical data. These situations occur frequently at the cutting edge of scientific work. In Collins' research on early gravitational wave detection, odd results could be attributed to a variety of sources, because there was, by definition, no universally accepted interpretation (Collins, 1975, 1981a, 1985). In particular, data conflicting with current theory could indicate either counterevidence to that theory or a flaw in experimental technique. The formation of scientific knowledge involves social interactions to reduce this variability in interpretation to the point that one conception wins out and becomes accepted as fact. This concept can also be applied to the development of technology. Different actors will have different conceptions of what existing technologies are, the nature of current problems, and what should be valued in potential future solutions (Pinch & Bijker, 1987).

With regard to both of these concepts, we argue that in inquiry in general, the role of argumentation is central (Bricker & Bell, 2008). The development of scientific and technological knowledge involves making substantive arguments using empirical and theoretical warrants. Indeed, we use this as a litmus test of inquiry instruction. It must require and enable students to make non-deterministic, empirically supported arguments at some point in the experience. By non-deterministic, we mean to exclude cases (sometimes found on standardized tests) where evidence points (and is often designed to point) in a clear, predictable direction.

Why is this hard?

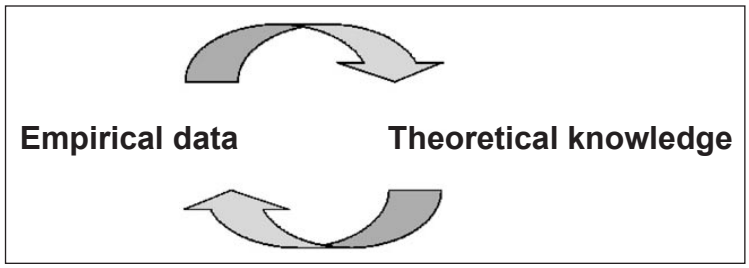
Creating the circumstances in which students can make these types of arguments often runs into two problems, which we will term the Getting on Board Problem and the Variability Problem.

The Getting on Board Problem can be illustrated by considering the simple diagram in Figure 1.

This is a very generic diagram of the development of knowledge in scientific communities. Phenomena are observed, and this generates the need to describe, organize, and explain it. This results in the development of new theoretical knowledge. This new framework suggests opportunities for new empirical observations. The new observations lead to new explanations, and so on. This is a cycle, and as member of a community, one never needs to consider a beginning. In the case of individuals, new participants begin by piggybacking on the work in progress of others, wherever in the cycle that might be. (For example, new Ph.D. students begin by working on the ongoing work of their mentors.) Once on board, they are part of the cycle, and are never working outside of a historical context.

This works in the general science community. However, in the classroom science setting, particularly at the pre-college level, things become problematic. Very few science classrooms are structured to allow junior students to benefit from interactions with more advanced students. Rather, students are grouped to all be at the same level.[1] Furthermore, students enter into scientific investigations with little or no background. Hence, no matter what part of the cycle we

Figure 1: A simple cycle of scientific work



1. We should note that typical heterogeneous classrooms do not do enough to remedy this problem. Mirroring science communities requires the interaction of individuals that are not just at different abilities but have different experiences, roles, and objectives. Having students with a variety of aptitudes but all of whom are engaging in a particular activity for the first time does not overcome the problem.

choose to inject students into, it is very difficult for it to have any meaning to them.

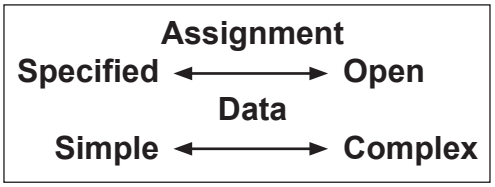
The Variability Problem stems from the need for a real argument. Making an argument means making statements about relationships. If students are only determining isolated attributes—say, the solubility of a particular chemical—there is no argument to have. Some sort of relationship among variables[2] is needed to create the tension that makes investigations meaningful. Furthermore, as noted above, we need the arguments to not be deterministic. This means having a degree of messiness to the data. There must be something there to argue over!

This requirement is also not easily achieved in the pre-college science classroom. First, the greater the technical resources of a classroom, the greater the opportunity to have data over which arguments can be had. This threshold can often be beyond the capabilities of classrooms. Second, the content most pre-college classrooms focus on is often very well established. The arguments have simply already happened. Lastly, understanding that there *is* tension or ambiguity in data generally depends on prior knowledge, leading us back to the Getting on Board Problem.

Balancing Acts

Designers of inquiry-based instruction can re-conceptualize these two problems as two balancing acts as shown in Figure 2.

Figure 2: The balancing acts of inquiry design



The first balancing act concerns the challenge or task given to students. This can range from very specific and rote to very open ended and ill-defined. Each end of the spectrum has problems. The specific end is the traditional cookbook lab, with all of its well-deserved criticism. There will be no variation in data (if done correctly). The goal is not to make a data-supported argument but to follow directions accurately in order to achieve the predetermined outcome (Amerine & Bilmes, 1990). There is nothing to argue over. But the open end of the spectrum also has problems. There is a limit on how open a task a given set of students can handle. The question is not merely whether or not the students can accomplish the task. It is conceivable that a task that students might technically not complete could still enable them to learn a tremendous amount through the effort. The problem arises when students are unable to determine how to make any movement whatsoever on a task. In other words, the task is so unintelligible to students that they cannot even proceed in a wrong direction, and therefore, also have nothing to argue over.

The role of the inquiry designer is to create a challenge/task/question that is understandable by the student as a challenge/task/question but not as a solution. Understanding an assignment as a solution means that there will be no argumentation over what is the proper solution. It is worth also noting that the right balancing point depends tremendously on the students. The challenge "build a car" is not an appropriate task for most middle school students. They wouldn't know where to start. However, this challenge might be appropriate for an advanced undergraduate mechanical engineering class. (There are other problems with this assignment that we will discuss a bit later.)

The second balancing act is on the data space that students will work with. Again, it is the extremes that are the problem, but also easy to fall into. The simplistic end is the traditional confirmation lab. The simple data is one-dimensional and conforms to the expected pattern, which students dutifully report. Since the data is straightforward, any arguments and their outcomes are predetermined.

The role of the inquiry designer is to create a challenge/task/question that is understandable by the student as a challenge/task/question but not as a solution.

Any deviation is automatically treated as a sign of poor experimental technique rather than possible support for alternative conclusions. On the complex end, you have data that is beyond the abilities of the students to collect and/or evaluate. This might be for conceptual reasons or for technical reasons. The data that can be collected with a particle collider offers plenty of opportunities for multiple interpretations—and therefore arguments—but such equipment is beyond most secondary schools!

So the task of the inquiry designer is to find (or create) a data space that

2. This does not need, however, to be quantitative in nature.

is approachable by students but also has some work for students to do. In other words, the data must be usable in terms of making arguments but not so usable that there is only one obvious argument available. Again, the right balance point depends on the particular students.

Frameworks

So what to do about these problems? We do not believe there is an easy, step-by-step way to produce inquiry instruction. Creativity and context will always be an essential element. However, in facilitating pre-service and in-service teachers' development of inquiry activities, we have noticed a pattern in instructional plans that seem to overcome the barriers we have discussed. We propose three models that inquiry designers can use to produce more inquiry instruction.

We must make two important caveats. First, we do not claim that these models are anything new. There are numerous examples of these models already in existence. Rather, we aim to put a label on them and, thus, identify how they are instances of a common phenomenon. By identifying a common pattern, we hope to provide guidance for generating new activities. Second, we do not claim that this list is exhaustive. There are certainly other sound inquiry-based activities that do not fall neatly within these forms.

Protocol Model

The Protocol Model has its origins in the Environmental Inquiry Project at Cornell University (ei.cornell.edu). A *protocol* is a well-defined procedure for collecting data. In terms of definition and clarity of steps, it is quite similar to a traditional cookbook lab. However, it is clearly portrayed as being just a tool—as opposed to the entirety of the lab experience. More importantly, a protocol can be applied to a wide variety of situations—not just the situation in which it is introduced and learned. (Hence, some cookbook labs can be adapted to form protocols but others cannot.) Once the students learn the protocol in an initial circumstance, they can then apply it to further research. This research can be more varied and more student-directed.

The prototypical case of a protocol is the lettuce seed bioassay (Trautmann, 2001a, 2001b). Students are given fairly clear directions for producing a serial dilution of a salt solution, setting up a bioassay using lettuce seeds, and evaluating the results. Once they have had that experience, they can now engage in further, more varied research: other concentration ranges, other toxins, and even other biological indicators. At the most sophisticated end of the spectrum, the bioassay can become a moderate piece in a larger extensive research endeavor.[3]

Another example of a protocol is the Watershed Habitat Evaluation and Biotic Integrity Protocol (WHEBIP) (Carlsen and Trautmann, 2004). This protocol was created to allow scientists to use models to predict aquatic biodiversity in watersheds. In this protocol, stream integrity ratings are assigned using land use criteria and can be accomplished using aerial photographs or remote sensing without requiring ground truthing (although in some instances, it is appropriate). Ratings are based on information students assess, including size of riparian belt, type of land use near stream, gradient, pollution, and conservation activity. Students can use this protocol to make a preliminary assessment of a habitat and, if desirable, make comparisons to data gleaned from ground truthing. This tool enables students to obtain data for one or multiple sites within watersheds or comparative studies between watersheds and make recommendations for remediation.

Learning a protocol is not just a question of now having a new technical skill. The student has also been introduced to a way of looking at the natural world. The dataset they produce in the initial learning round is also significant. It can be an indicator of what aspects of the phenomenon merits investigation next, just as with science at large. Hence, the student has been brought on board the knowledge development cycle.

A counter example can help illuminate the nature of an effective protocol. A common physics cookbook lab is to measure the period of a pendulum with various lengths and masses. Unlike some cookbook labs, this is not easily configured into a protocol. It fails to overcome both problems. The data produced is not likely to have any ambiguity—and any that does occur will be attributed to practitioner error. In addition, once the initial data is collected, then what? The experience will not introduce students to a new empirical realm.

Design Challenge Model

The Design Challenge Model has had more common use. Design Challenges are centered on an explicit

3. This is one way we have seen where (with a lot of work) the Getting-on-Board problem can be overcome. Seniors carrying out an extensive research project "contract out" their bioassay needs to lower grade students, providing them the opportunity to learn the basic procedures (Avery, 2003).

production task. Often the task will motivate the practical need to acquire certain knowledge bases. Sometimes inquiry designers will use a jigsaw arrangement in which students are divided into specialty groups to learn one of the applicable knowledge bases, then rearranged into design teams made up of representatives from each specialty group.

Forming the explicit charge that is given to the students is the critical and creative focal point of designing Design Challenges. Accomplishing this goal can determine if the balancing acts have been achieved. As mentioned above, a way of framing the problem is to give students a challenge that is understood as a question but not as a solution. A question for which students already have a single, preconceived solution will not generate the argument opportunity necessary for inquiry. At the other end of the spectrum, a task for which students have no conception or ability to proceed is equally unfruitful. However, it should also be noted that there is another way in which Design Challenges can be too open-ended. Consider the challenge for middle school students "design and build a paper airplane." This avoids both of the problems noted so far. Students understand what a paper airplane is, they have the intellectual and material resources to meet this challenge, and they are likely to propose multiple solutions. But then what? The litmus test we described requires *warrant*-based arguments—essentially saying "this is better because of such and such". As stated, this design challenge does not include any means to defend why one design is better than another. The challenge is *too* open—not in term of students' cognitive or technical ability to achieve it, but in terms of it being a meaningful competition. Design Challenges must include pressures that require student designers to make judgments and back up those judgments with arguments. Hence, the "build a car" challenge noted above would be an ineffective design challenge even for a group of students that could build a car. More pressure is needed.

We do not believe there is an easy, step-by-step way to produce inquiry instruction.

Design challenges often result in tangible products. One example of this is the stormwater treatment design challenge (Carlsen and Trautmann, 2004). This activity models how cities develop systems for collecting and draining runoff from storms. Using simple materials such as plastic soda bottles, tape, coffee filters, cat litter, sand, gravel, and plastic tubing, students are given the task of creating a filtering system that can handle a simulated storm event over a relative period of time. They need to take into account the various types of substances (such as chemicals, dirt, oils, etc) found in runoff, the volume of the storm event, the time between events and the extent to which the stormwater needs to be filtered. Like engineers in the real world, they are also constrained by materials, guidelines, budget, time, and design. From a curriculum design point of view, the specifics of the design constraints (size, materials, etc.), evaluation measures (ph, DO, etc.) and simulated runoff (particulate matter, oil, etc) will be what determine how the balancing acts have been achieved.

Just making something, however, does not make an effective design challenge. Construction activities can be the Design Challenge equivalent of a cookbook lab. Consider the common example of students in physics classes designing roller coaster rides. The details of the assignment are crucial in determining whether this is an effective Design Challenge. Often, students design the ride in a fairly arbitrary way, and then post facto apply physics principles to determine elements like speed. The laws of physics do provide limitations on the design (e.g. a hill can not be too high that a car will not have the energy to reach the top), but there are not competing constraints that provide for points of debate. Once a student stays within the bounds of physics, any choice is an arbitrary preference, and hence, there is no opportunity for argumentation.[4] This illustrates how the details of an assignment can have a profound effect.

Although producing a tangible object is perhaps the most common type of design challenge, we should not limit our students or our own thinking to this format. Consider the following example (Meyer, 2003). Students are given a scenario in which a community that is experiencing pollution in a local waterway. The students are divided into different constituency groups: farmers, homeowners, industry, and municipal authorities. They are given a variety of information resources—some common and some specific. They then

4. We should note that we are not arguing that such an activity is not worthwhile, but simple that it does not work as an effective design challenge.

have a variety of meetings—some in homogenous groups and some in heterogeneous groups. The task and final outcome of those meetings is to develop a restoration plan. This example creates opportunities for debate without being too open-ended, but it results in a plan of action rather than a physical product.

Product Testing Model

In general, the Protocol Model and Design Challenge Model can be seen as corresponding to scientific work and engineering work respectively. We have used these frameworks with pre-service and in-service science teachers and feel they genuinely represent general frameworks that can be utilized to inspire and guide further design of inquiry instruction. We end by proposing a third framework. It will take further work to flesh out its details and legitimacy.

The Product Testing Model is inspired in part by the Discovery Channel show Mythbusters (Rees, 2003). A common thread through much of the work on the show, and product testing in general, is the challenge to reproduce natural phenomena under lab conditions—i.e. in an intentional, controllable, measurable, and reproducible manner. In this sense, it is much like a sub-set of design challenges. But there is also a second point of contention: once results are obtained, how should they be evaluated? Consider the task of determining the best paper towel. What makes the best paper towel? How can a desired characteristic like durability be measured (in order to make clear that brand A is more durable than brand B)? And once that is done, how should durability be related to other characteristics, like price? Hence the Product Testing Model operates in two problem spaces: physically performing the relevant tests and determining criteria for success and failure. In a way, it is the combination of the Protocol and Design Challenge Models. A task generates the needs for various knowledge domains and the development of data collection routines.

Conclusion

We have put forward a framework that identifies two critical problems in designing inquiry-based instruction and suggests three models for developing instruction that overcomes those problems. The Protocol Model overcomes the Getting on Board Problem by providing students an initial experience through clearly delineated steps with a data collection technique that can be applied to a wide variety of further settings. It not only gives students a new tool, but also suggests questions to which it can be applied. It overcomes the Variability Problem by being applicable to a wide variety of settings and utilizing messy data. The Design Challenge Model overcomes the Getting on Board Problem by presenting a practical need to acquire certain knowledge bases. It asks students to understand it as a question before understanding a solution. It overcomes the Variability Problem by imposing a variety of pressures that allow students to balance competing needs in a variety of ways. Lastly, the Product Testing Model overcomes the Getting on Board Problem by centering on everyday phenomena. It overcomes the Variability Problem both through the challenge of producing the phenomena in the lab setting and through competing values.

References

Amerine, R., & Bilmes, J. (1990). Following instructions. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice* (pp. 323-336). Cambridge, Mass.: The MIT Press.

Anderson, R. D. (2002). Reforming Science Teaching: What Research says about Inquiry. *Journal of Science Teacher Education, 13*(1), 1-12.

Avery, L. M. (2003). Knowledge, Identity, and Teachers' Communities of Practice. Unpublished Doctoral Dissertation, Cornell University, Ithaca.

Beerer, K. and Bodzin, A. (2003). Promoting inquiry-based science instruction: The validation of the Science Teacher Inquiry Rubric (STIR). Journal of Elementary Science Education, 15(2), 39-49.

Bricker, L. A., & Bell, P. (2008). Conceptualizations of Argumentation From Science Studies and the Learning Sciences and Their Implications for the Practices of Science. *Science Education, 92*, 473-498.

Carlson, W.S., & Trautmann, N.M. (2004), *Watershed Dynamics*, Arlington, VA: NSTA Press.

Collins, H. M. (1975). The seven sexes: A study in the sociology of a phenomenon, or the replication of experiments in physics. *Sociology, 9*, 205-224.

Collins, H. M. (1981a). Son of seven sexes: The social destruction of a physical phenomenon. *Social Studies of Science, 11*, 33-62.

Collins, H. M. (1981b). Stages in the empirical programme of relativism. *Social Studies of Science, 11*, 3-10.

Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. Chicago: The University of Chicago Press.

Herron, M.D. (1971). The nature of scientific inquiry. *School Review, 79,* 171-212.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Chicago: University of Chicago.

Meyer, D. Z. (2003). Social Engagement in Curriculum Design. Unpublished Doctoral Dissertation, Cornell University, Ithaca, NY.Minstrell, J., & van Zee, E. (Eds.). (2000). *Inquiring into inquiry learning and teaching in science*. Washington, D.C.: American Association for the Advancement of Science.

National Research Council. (1996). *National science education standards*. Washington: National Academy Press.

National Research Council. (2000). *Inquiry and the National Science Education Standards*. Washington, DC: National Academy Press.

Pinch, T. (1981). The Sun-set: The presentation of certainty in scientific life. *Social Studies of Science, 11*, 131-158.

Pinch, T. (1985). Towards an analysis of scientific observation: The externality and evidential significance of observational reports in physics. *Social Studies of Science, 15*, 3-36.

Pinch, T., & Bijker, W. (1987). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. In W. E. Bijker, T. P. Hughes & T. J. Pinch (Eds.), *The social construction of technological systems* (pp. 17-50). Cambridge, Mass: MIT Press.

Rees, P. (Writer) (2003). Mythbusters: Discovery Channel.

Trautmann, N. M. (2001a). *Assessing Toxic Risk: Student Edition*. Arlington, VA: NSTA Press.

Trautmann, N. M. (2001b). *Assessing Toxic Risk: Teacher's Guide*. Arlington, VA: NSTA Press.

Wheeler, G. F. (2000). The three faces of inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 14-19). Washington, DC: American Association for the Advancement of Science.

Daniel Meyers is an assistant professor, Department of Mathematics and Science Education, Illinois Institute of Technology, Chicago, IL 60616. Correspondence concerning this article may be sent to <meyerd@iit.edu>.

Leanne M. Avery is associate professor of science education, Department of Elementary Education and Reading, State University of New York College at Oneonta.

Patrick M. Lloyd, Ronald A. Eckhardt

Strategies for Improving Retention of Community College Students In the Sciences

Almost one half of U.S. students receiving B.S. and M.S. science degrees attend community colleges during their academic careers, yet for the large majority of community college students in the sciences, a four-year degree in a STEM discipline remains an unrealized goal. The authors describe methods intended to improve student learning, retention, and graduation rates of community college students in the sciences.

Introduction

Community colleges play an important role in the education and training of students in the sciences. In 2004, nearly half of the Bachelor's and Master's degrees awarded in the United States in science and engineering were granted to students who had attended community colleges at some point during their academic careers (Kincaid, et al., 2006; Tsapogas, 2004; Ryan, Wesemann, Boese, & Neuschatz, 2003). The role of these two-year institutions in science education has not been overlooked by policy-makers. The National Science Board has identified the importance of community colleges in developing a technical workforce that can allow United States companies to compete with their Chinese and Indian counterparts (National Science Board, 2006; U.S. Dept. of Education, 2000). The National Institute of Health funds a Bridge program that targets minority science students for transfer from community colleges into baccalaureate programs (Carpenter, 2008). Although there is widespread acknowledgment of the importance of community colleges in training workers, educators have made limited progress in helping the majority of students at community colleges to reach the levels of academic and economic success enjoyed by their counterparts at four-year colleges and universities. In the sciences, community colleges award associate's degrees in fields where bachelor's, master's, and doctorate degrees are increasingly becoming a requirement for employment (National Academies Press, 2007). This makes facilitating transfers from community colleges to four-year colleges an essential goal for educators and administrators.

Although there is widespread acknowledgment of the importance of community colleges in training workers, educators have made limited progress in helping the majority of students at community colleges to reach the levels of academic and economic success enjoyed by their counterparts at four-year colleges and universities.

Socioeconomics can play a role in determining which students attend community colleges. Students enrolled in community colleges tend to be financially disadvantaged compared to students who enroll in four-year colleges directly out of high school (Government Accounting Office, 2008). High school graduates who receive diplomas with college preparatory courses and have the financial ability generally progress directly to four-year colleges. Students without the requisite high school preparation or financial ability often

attend the more affordable community colleges (Phillippe & Sullivan, 2005).

Many students who begin their college studies at community colleges intend to graduate quickly and move on to more advanced degree programs (Rouse, 1999; Leigh & Gill, 2003). Despite the best intentions of community college faculty, most students leave the community colleges without obtaining degrees or transferring to four-year colleges. Estimates suggest that transfer rates for students from community colleges to four-year colleges can be as low as twenty percent for students wishing to do so (Bradburn & Hurst, 2001; Gordon, 1996). Likewise, graduation rates at community colleges are as low as thirty percent (Wild & Ebbers, 2002; Mohammadi, 1994). Because many students start their science and engineering careers at community colleges, it seems worthwhile to develop and implement strategies that improve the effectiveness of science education at these institutions in order to better prepare students to transfer to and succeed at four-year colleges.

Recent publications have studied institutional policies that affect the success rates of community colleges in preparing students for more advanced academic work (Striplin, 1999; Cohen & Brawer, 1996). Recommendations include institutional changes that affect the methods used to fund community colleges and improving counseling and advising services (Burgess & Samuels, 1999; Cohen & Brawer, 1996). It is also worthwhile to consider student performance following completion of the developmental courses and entry into college-level science courses (Long & Kurlaender, 2008). In this report, we discuss the application of methods for improving student success rates in the first two years of science education through a program called the Brooklyn Gateway. Specifically, we focus on improving student performance in a freshman-level general chemistry course. Each method is relatively inexpensive to implement. Taken as a whole, they require coordination among educators within the science curriculum. It is also necessary to consider the realities of community college students' lives, because these realities affect student utilization of support methods and student response to format of instruction.

Despite the best intentions of community college faculty, most students leave the community colleges without obtaining degrees or transferring to four-year colleges.

General Chemistry and Student Success in STEM Majors

Our institution's student retention rate is comparable to many urban community colleges. The college-wide six-year graduation and transfer rate to four-year institutions are each around thirty percent. For science and engineering majors, retention and transfer rates are similar to those of other academic areas at the College. However, unlike students in the humanities, important challenges to the success of science and engineering students include mathematics and science courses that require a high proficiency in mathematics. At our institution, only science and mathematics majors are required to take college-level math courses. Biology and engineering science are the two largest programs in the sciences at our institution. In reviewing graduation rates in these two programs, we identified general chemistry as a particularly significant stumbling block for students. Historically, pass rates in general chemistry have hovered near fifty percent. Our plan was to improve student performance in this important gateway course so that students would be more likely to progress through their academic programs, including pursuit of more advanced courses within their disciplines.

College-level general chemistry is, on the whole, a difficult course for many science students, including students at four-year colleges (Chambers, 2005). For many community college students, general chemistry is particularly problematic. Administrators at our institution have referred to general chemistry as a "killer course." Counselors often recommend to students that they avoid the course until their last year of school so that their grade point average won't be dramatically affected (Phillip, Brennan, & Meleties, 2005).

Preparation is a key component to success. In New York State, the majority of community college students have not completed a one-year course in high school chemistry or physics. Graduation requirements for high school students include completion of two Regents' level science courses (New York State Education Department, 2009). Although students must enroll in science courses that satisfy state graduation requirements, they are not required to complete science courses designed as college preparatory courses (Haycock, 2001; Gamoran, 1987). Consequently, many students

graduate high school without the skills necessary to succeed in science and engineering disciplines. It is only when students find themselves in college that they realize that allied health fields and engineering disciplines require challenging survey courses like general chemistry and general physics.

Many students are turned off by the perception that chemistry is not a subject pertinent to their career goals.

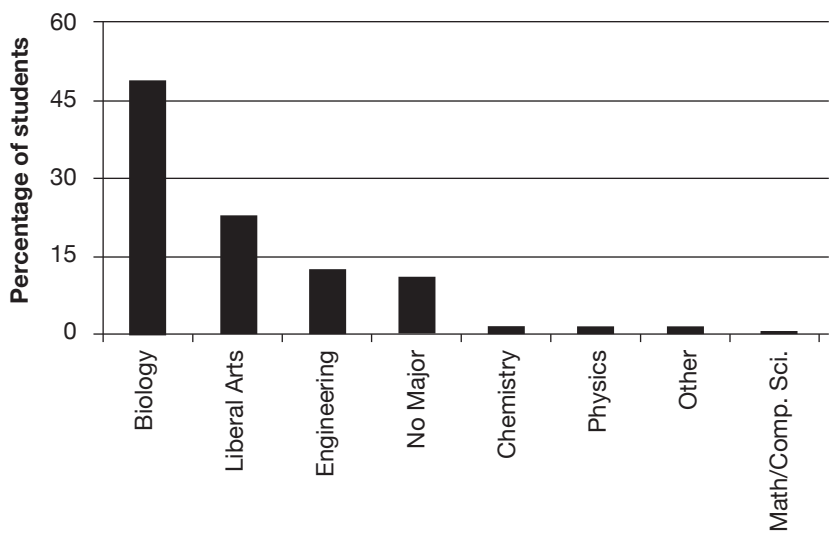
At two-year institutions, general chemistry is often the first science course required of students that involves a high level of quantitative reasoning skills. It is also often the first course in which students must connect physical theories with specific sets of calculations. The application of graphical analysis to experimental data and extraction of physical parameters are also challenging objectives for many students. These tasks are often overwhelming for students with minimal high school backgrounds in math and science. To prevent under-prepared students from enrolling, chemistry departments have generally instituted math pre-requisites for general chemistry. Others have developed preparatory courses that focus on developmental topics like scientific notation, significant figures, factor-label analysis, and rudimentary chemistry skills like equation balancing and chemical nomenclature. The effectiveness of these strategies has been the subject of some discussion (Bentley & Gellene, 2005; Jones & Gellene, 2005). Our experience suggests that preparation in quantitative reasoning is not the only, or even most important, stumbling block for students. Some of the concepts that students in freshman-level chemistry find the most difficult do not involve intensive calculations. Examples include net ionic equations, quantum chemistry, and bonding theory. Nearly all of the students at our institution enrolled in general chemistry have completed the math requirement for the course. Many have received high marks in their math courses and passed the preparatory course, but still fare miserably in general chemistry.

Another challenge is student motivation. Many students are turned off by the perception that chemistry is not a subject pertinent to their career goals. Unlike four-year institutions, many science majors at community colleges are not studying to enter medical or pharmacy programs. Most biology majors at our institution are interested in allied health professions like physical therapy, physician assistant, and nursing (Figure 1). Abstract concepts like electron configurations and orbital hybridization models seem irrelevant to many students. This perceived disconnectedness of subject matter to professional goals leads to low morale in the course and, consequently, lower student performance (Gillespie, 1997).

Methods for Improving Instruction

Our strategy to improve student learning and retention in general chemistry was to find a way for students to dedicate as much attention to the course as possible and to provide the support they needed to do so. The regular twelve-week semester is problematic for many students. They often enroll in four or five courses and have little time to focus on any of their classes. It is not uncommon for students to begin the semester with five courses and finish with two or three as they withdraw from the more challenging and time-consuming courses. For that reason, we enrolled students in an immersion chemistry section during the shorter six-week sessions the College offers each summer and winter (Table 1). In these

Figure 1: A breakdown of majors enrolled in general chemistry



It is not uncommon for students to begin the semester with five courses and finish with two or three as they withdraw from the more challenging and time-consuming courses.

sections, the preparatory chemistry pre-requisite was waived. Immersion sections were enriched compared to traditional six-week summer and winter sessions by adding Peer-Led Team Learning (PLTL) sessions twice a week, daily optional drop-in tutoring, and group trips to science learning centers on off-days (Stewart, Amar, & Bruce, 2007; Gosser & Roth, 1998; Woodward, Gosser, & Weiner, 1993). Immersion sections were taught by faculty members who limited their research activities for the duration of the course in order to increase student access and foster instructor-student mentoring. Students received a three hundred dollar stipend for participation in the course as compensation for the cost of textbooks and supplies and the additional time spent on the course.

Table 1: Comparison of modes of instruction in general chemistry

	Lecture	Laboratory	PLTL Workshops	Drop-in Tutoring
12-week (traditional)	4 hours	2 hours	N/A	N/A
6-week (traditional)	8 hours	4 hours	N/A	N/A
6-week (immersion)	8 hours	4 hours	4 hours	Daily

Student Performance in Accelerated and Immersion Sections

General chemistry courses at our institution are taught by both full-time tenure-track faculty members and part-time contingent instructors. Enrollment is capped at twenty-five students per section with typical enrollments of around twenty students. We report here the results for six immersion groups for the time period between 2006 and 2008 (123 students). We offer as a comparison the students who enrolled in the regular twelve-week semester for the time period between 2001 and 2008 (1389 students) and students who enrolled in the traditional six-week semester between 2001 and 2008 (548 students).

We observed that there is a small difference in the percentage of students who received a grade of C or higher for those who enrolled in the traditional six-week sessions compared to the traditional full-length twelve-week sessions (Figure 2). Students enrolled in the shorter sessions generally have a slightly higher pass rate (C or better) than the students enrolled in the twelve-week sessions (55% compared with 50%). We also calculated a numerical score for students on a traditional four-point scale (Table 2). We then calculated an average score for each section of students which we call the course average.

The course average was found to be somewhat higher for students enrolled in the shorter six-week sessions (1.73 versus 1.46) (Figure 3). This difference is consistent with grade distributions in other chemistry and science courses, and we believe that it is a reflection of how college fits into our students' lives. The majority of students commute by way of public transportation and one-way travel times are typically one to two hours. Students often hold part-time or even full-time jobs in addition to attending college. A sizable number of students are parents or care for family members. Financial aid considerations weigh heavily on the students' academic schedules. Students are required to make regular academic progress in order to maintain access to student aid and public assistance funds. Aid agencies determine how many and into which courses students may enroll. During the six-week sessions, students are limited to enrollment in a maximum of two courses and typically enroll in one science course and one humanities course. The duration of these off-sequence courses is half that of courses during the regular twelve-week semester. Despite this acceleration, students may have a better ability to focus on their studies because they enroll in fewer courses. Another advantage of the accelerated schedule may be that faculty members are better able to dedicate time to the course and to students, because they generally teach only one section compared to three or more sections during the twelve-week sessions.

Table 2: Course scoring scale

Course grade	Point Value
A	4.0
B	3.0
C	2.0
D	1.0
F or Withdraw	0.0

The difference in academic performance of students enrolled in accelerated sections compared to twelve-week sections was one reason for considering immersion sections during the winter and summer sessions. We thought that we might be

able to take advantage of the course-load limit placed on students in the shorter sessions. Improvements in the distribution of course grades were apparent when results for the first immersion group were tabulated. In the first immersion section fourteen of the fifteen students enrolled completed the course. Of those fourteen students who completed the course, all but one received a passing grade. Students in each of the subsequent immersion sections succeeded in the course with higher than normal pass rates.

In the six immersion sections studied, students completed the course with higher grades in the course than those in the twelve-week sections as well as the traditional six-week sessions (Figures 2 and 3). The average pass rate for students in six immersion groups was

81%, and the cumulative course grade-point average was 2.43. Students in the immersion sections performed better as a group than students in traditional sections offered simultaneously in each of the six sessions studied.

The percentage of students receiving a letter grade A was not significantly higher for students in the immersion groups compared to students in traditional sections. This suggests to us that the program has had a particular effect on students at risk of failing the course. In exit interviews, many students stated that the course was more difficult than they had expected. Some were aware that the "extra" support had an effect on their course outcomes. Some students offered the criticism that the support made the course time-consuming and challenging, even though they performed better than students in traditional sections. It seems possible that many students may be unaware of the time commitment necessary to succeed in the course and, consequently, become overwhelmed when taking four or five courses.

Organic Chemistry as an Indication of Progress

The addition of Peer-Led Team Learning sessions, after-class tutoring, and greater faculty involvement may have helped bring about increases in course pass rates and grades-point averages. An important question is whether some of the positive effects of the program continue in future courses (Horwitz, & Rodger, 2009; Becvar, Dreyfuss, Flores, & Dickson, 2008). To study this possibility, we looked at the organic chemistry course. Organic chemistry is a sophomore-level course for chemistry, pharmacy, and pre-med students offered by our department. Enrollment in the course includes a mix of students who completed general chemistry at our institution and students from other institutions. To determine if there was a benefit to students from the immersion program in organic chemistry, we compared students who successfully completed general chemistry in one of the first four immersion sections to those who completed the course in a traditional six-week section. The number of students in each group is similar with 73 students from the immersion sections and 74 from the traditional sections (Figure 4). The success rate in organic chemistry was higher for students from the immersion sections, with fifteen immersion students passing compared to eight from the traditional sections. The number of students from immersion sections that failed organic chemistry over the same period was two, compared to five students from the traditional sections. Although the numbers being considered are small, they do suggest that, for our students, organic chemistry is not a fundamentally more challenging course than general chemistry. In fact, failure rates in organic chemistry are significantly lower at our institution than are those in general chemistry. The pedagogical approach in organic chemistry is quite different from that of general chemistry. Although organic chemistry

Figure 2: Pass rates in general chemistry

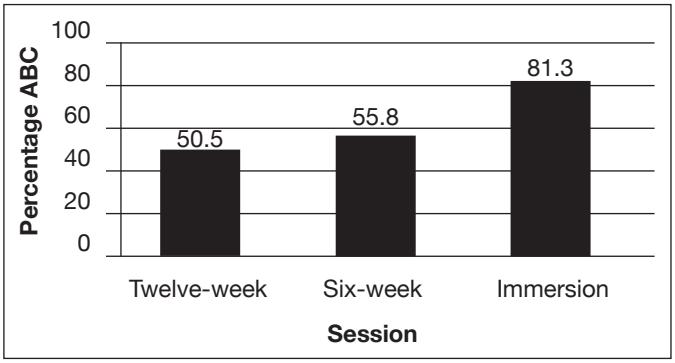
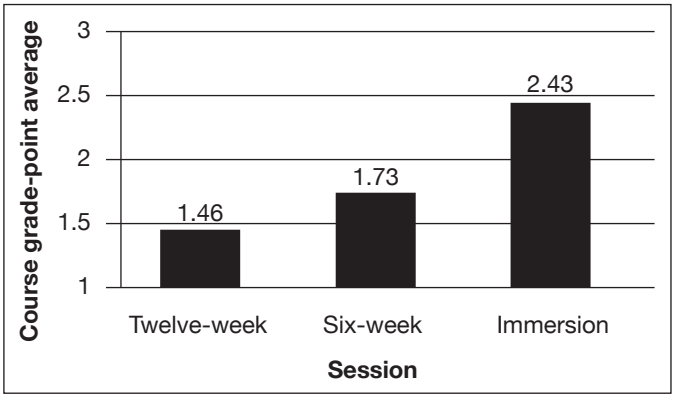


Figure 3: Course numerical average in general chemistry



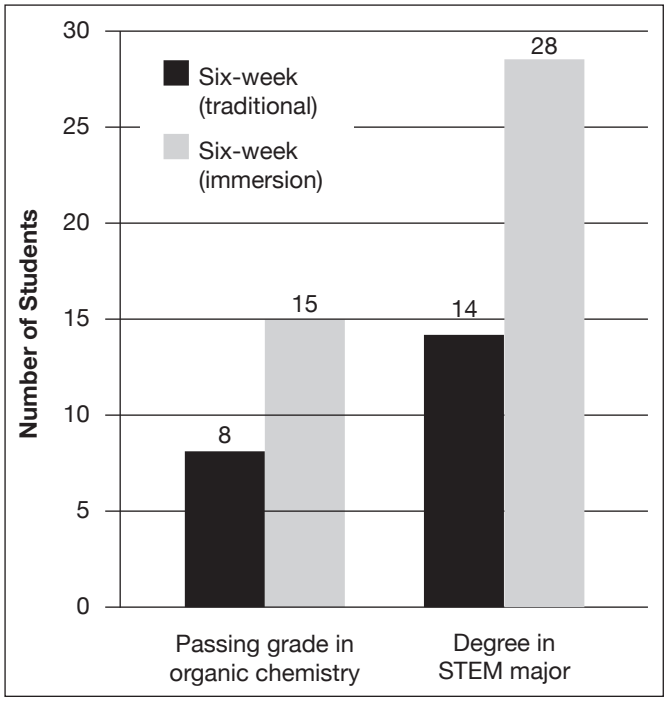
is considered a more advanced course, the two courses possess significant conceptual overlap. The focus on calculations in general chemistry often works to filter out students who may have many of the skills needed to advance but aren't proficient in numerical and algebraic calculations. It may be that by creating a mechanism for students to succeed in general chemistry, we have opened up a potential for students to succeed in more advanced chemistry courses. We are monitoring the progress of subsequent immersions groups as well as students who return to the College to complete organic chemistry after transferring to other institutions in order to determine if this pattern continues.

Increasing Graduation Rates in the Sciences

The three-year graduation rate at our institution has fluctuated between fifteen and twenty-five percent over the course of the past decade (CUNY Office of Institutional Research, 2005). Figure 4 shows the number of STEM degrees (science, technology, engineering, and mathematics) awarded to students who attended one of the first four immersion sections between 2006 and 2007, along with the number of graduates who attended traditional six-week sections during the same time period. The number of students initially enrolled in each of the two groups is similar, but the attainment of STEM degrees is twice as high in the immersion group (38% versus 19%). This increased graduation rate can be explained by considering the distribution of majors in our general chemistry course. Biology majors comprise the predominant group of students in our general chemistry sections (Figure 1). General chemistry is the terminal chemistry requirement at our institution for biology majors, and it is often considered by students to be their most challenging course. By increasing the pass rates in general chemistry, we may have eliminated the major barrier to graduation for the majority of those students (Reingold, 2001). We are monitoring student retention and graduation rates to determine whether the long-term number of graduates increases.

It may be that by creating a mechanism for students to succeed in general chemistry, we have opened up a potential for students to succeed in more advanced chemistry courses.

Figure 4: Pass rates in organic chemistry and achievement of STEM degrees within two years for students in immersion sections and traditional six-week sections



Peer-Led Team Learning

Increases in course grades and graduation rates are encouraging, and this suggests that there are students who may progress academically if we use some well-known methods for improving instruction. For example, PLTL is a method designed to introduce constructivist approaches to science education and has been used in the physical sciences since the 1990's (Fosnot, 2005; Vykotsky, 1978). In PLTL, a student peer who has previously succeeded in the course leads a group of six to eight students through faculty-designed workshops. Workshops are designed to promote exploration of the course material outside the traditional lecture environment. In PLTL workshops, the course instructor is absent as are answer keys to the workshop materials. Important goals include shifting the focus of education away from lecturing and toward active student learning, developing student leadership skills, and democratizing learning. We used PLTL as a method of instruction in order to help students develop a deeper understanding of the material and as a way to discourage memorizing algorithms as a method of solving problems. These are particularly important goals, because many of the students transferring from our institution suffer from a "transfer shock" when they

arrive at four-year institutions (Diaz, 1992; Cejda, 1997). Class sizes in the sciences are often much larger at senior colleges. Students there are often more competitive compared with students in community colleges. Faculty members are also often less available to students than those at community colleges. These differences are factors in the decisions of large numbers of students to leave the sciences shortly after reaching the four-year colleges. One role of PLTL is to help students to develop their own reasoning and critical-thinking skills through practice in a social environment and to promote a sense of independence that allows them to be less dependent on instructors.

The Role of Tutoring

The number of students who used after-class tutoring fluctuated between four and eight students out of twenty in each of the immersion sections. Although these numbers seem low, they are higher than we have observed for students in traditional tutoring services offered by the College. Because many of our students work after class, they were often unable to take advantage of tutoring. However, the number of students who used tutoring spiked near midterm and final examinations. This suggests that providing flexible student support is an important component in connecting students to support. The amount of academic support students receive outside the classroom at our institution is limited. Traditional tutoring at the institution takes place through semester-long scheduled appointments. If students miss a total of two sessions for any reason, they are dropped from tutoring for the remainder of the semester. Conversely, during the immersion sessions, we offered more flexible, drop-in tutoring, and this was found to be an important contribution to the program. The highest attendance in tutoring sessions was found when tutors also served as peer leaders from the PLTL workshops. We believe that student attendance in tutoring sessions may have acted as a measure of their confidence in the tutors. Successful workshops often led to working relationships between leaders and student participants that expanded beyond the time constraints of the workshops and into tutoring sessions. We also offered drop-in tutoring to students who were in the traditional sections. With few exceptions, those students did not make use of tutoring. We believe that this shows that tutoring is more important to students when it is connected to other components of the course.

Student-Faculty Interactions

As part of the immersion program, we invited students to join their instructors at science institutions like the American Museum of Natural History and the New York City Hall of Science. Attendance among students ranged between twenty-five percent and fifty percent. Just as in tutoring, one of the challenges is the large number of students who work during the times we were able to schedule the trips. For the students who did attend, there was a tendency for them to become more connected to the other students in the course, peer leaders, and faculty members. Some students became peer leaders in future PLTL sessions, leading to continuity of the program. A few others became involved in laboratory research projects with faculty members, leading to even greater connection to their academic disciplines (Gafney & Varma-Nelson, 2007).

Future Directions

Based on the successes of the immersion program, we have implemented some of the same methods for the twelve-week semester. Students were enrolled in sections that included PLTL workshops and after-class drop-in tutoring. We found a slight improvement in the course pass rates and grade distributions compared to historical groups, with pass rates slightly above sixty percent. However, these increases in student retention and grade performance were lower than those observed for students in the six-week immersion groups. We believe that enrollment in four to five courses per semester is a challenge for many students due to the realities of their lives. One strategy we are considering is a schedule that involves two six-week chemistry courses, taken serially during the traditional twelve-week semester. Under this plan, students would be limited to enrolling in one non-chemistry course during the twelve-week semester.

Important goals include shifting the focus of education away from lecturing and toward active student learning, developing student leadership skills, and democratizing learning.

We have yet to determine the effectiveness of the immersion program on student academic performance following transfer to four-year institutions. We are monitoring the academic progress of students who have transferred to local four-year public colleges and have

implemented PLTL in two courses at a local four-year sister campus. We chose advanced physiology as well as first-semester organic chemistry, because these are courses in which biology majors commonly enroll during the transition from their sophomore to junior year. Our goal is to provide a bridge for students once they enter a four-year program by promoting learning environments that are similar to those they experienced during their years in community college.

References

Becvar, J.E., Dreyfuss, A.E., Flores, B.C., & Dickson, W.E. (2008). 'Plus-two' Peer-led team learning improves student success, retention, and timely graduation. *38th ASEE/IEEE Frontiers in Education Conference*.

Bentley, A.B. & Gellene, G.I. 2005). A six-year study of the effects of a remedial course in the chemistry curriculum. *Journal of Chemical Education*, 82, (1), 125-130.

Bradburn, E.M., & Hurst, D.G. (2001). Community college transfer rates to four-year institutions using alternative definitions of transfer. *Education Statistics Quarterly*, 3,3, 119-125.

Burgess, L. & Samuels, C. (1999). Impact of full-time versus part-time instructor status on college student retention and academic performance in sequential courses. *Community College Journal of Research and Practice*. 23, 487-498.

Carpenter, S. (2008). Community colleges fuel science workforce. *Science Careers*. DOI: 10.1126/science.caredit.a0800056.

Cejda, B.D. (1997). An examination of transfer shock in academic disciplines. *Community College Journal of Research and Practice*, 21, (3), 279-288.

Chambers, K.A. (2005). Improving performance in first year chemistry. Doctoral Dissertation, Texas Tech University.

Cohen, A. M., & Brawer, F. B. (1996). Policies and programs that affect transfer. Washington, DC. *American Council on Education*. (ED 385 336)

CUNY Office of Institutional Research and Assessment. (2005). Available at <http://oira.cuny.edu:7778/portal/page/portal/oira/CUNY_Data_Book_by_Table_Number_Fall_2004/trends.ret.grad.rates.system.ftf.assoc.pdf>.

Diaz, P.E. (1992). Effects of transfer on academic performance of community college students at the four-year institution. *Community College Journal of Research and Practice*, 16, (3), 279-291.

Fosnot, C.T. (2005). *Constructivism Theory, Perspective, and Practice, 2nd edition*. New York, NY. Teachers College Press.

Gafney, L. & Varma-Nelson, P. (2007). Evaluating peer-led team learning: A study of long-term effects on former workshop peer leaders. *Journal of Chemical Education*, 84, (3), 535-539.

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education*, 60, (3), 135-155.

Gillespie, R. (1997). Reforming the general chemistry textbook. *Journal of Chemical Education*, 74, (5), 484-485.

Gordon, J.A. (1996). Differential advantages in an innovative community college setting. (ED 397 883).

Gosser, D.K., & Roth, V. (1998). The workshop project: peer-led team learning. *Journal of Chemical Education*, 75, (2), 185-187.

Government Accounting Office. (2008). Community colleges and one-stop centers collaborate to meet 21st century workforce needs. Report to Congressional Requestors, GAO-08-547.

Haycock, K. (2001). *Educational Leadership*. Closing the achievement gap. 58, 6.

Horwitz, S. & Rodger, S.H. (2009). Using peer-led team learning to increase participation and success of under-represented groups in introductory computer science. *Proceedings of the 40th ACM technical symposium on computer science education*. 163-167.

Jones, B. and Gellene, G.I., (2005). Understanding attrition in an introductory chemistry sequence following successful completion of a remedial course. *Journal of Chemical Education*. 82, (8), 1241-1245.

Kincaid, W.B., Narum J., Koupelis, T., Shriner, W., Adams-Curtis, L., Agwu, N., et al. (2006). Bringing community college faculty to the table to improve science education for all. Project Kaleidoscope.

Leigh, D.E. & Gill, A.M. (2003). Do community colleges really divert students from earning bachelor's degrees? *Economics of Educational Review*. 22, (1), 23-30.

Long, B.T. and Kurlaender, M. (2008). Do community colleges provide a viable pathway to a baccalaureate degree? *Educational Evaluation and Policy Analysis*, 31, (1). 30-53.

Mohammadi, J. (1994). Exploring retention and attrition in a two-year public community college. (ED 382 257)

National Academies Press. (2007). Rising above the gathering storm: Energizing and employing America for a brighter economic future.

National Science Board. (2006). Science and Engineering Indicators. National Science Foundation. 1, NSB 06-01.

New York State Education Department. (2009). <http://www.emsc.nysed.gov/part100/pages/1005>.

Phillip, N., Brennan, T., & Meleties, P. (2005). Increasing enrollment in chemistry courses and programs. *The Chemical Educator*. 10, (3), 227-230.

Phillippe, K. & Sullivan, L. (2005). National profile of community colleges, 4th edition. American Association of Community Colleges. Reingold, I.D. (2001).

Bioorganic First: A new model for the college chemistry curriculum. *Journal of Chemical Education*. 78, (7), 869-871.

Rouse, C.E. (1999). Do two-year colleges increase overall educational attainment? Evidence from the states. *Journal of Policy Analysis and Management*. 17, (4), 595-620.

Ryan, M.A., Wesemann, J.L., Boese, J.M., & Neuschatz, M. (2003). The status of chemistry in two-year colleges: results from a survey of chemistry departments. American Chemical Society.

Striplin, J.J. (1999). Facilitating transfer for first-generation community college students. ERIC Clearinghouse for Community Colleges. EDO-JC-9905

Stewart, B. N., Amar, F. G., & Bruce, M. R. (2007). Challenges and rewards of offering peer-led team learning (PLTL) in a large general chemistry course. *Australian Journal of Education in Chemistry*, 67, 31-36.

Tsapogas, J. (2004) The role of the community colleges in the education of recent science and engineering graduates. *Science Resource Statistics*. NSF 04-315. National Science Foundation.

U.S. Department of Education. (2000). Land of plenty: Diversity as America's competitive edge in science, engineering and technology. Report of the congressional commission of the advancement of women and minorities in science, engineering, and technology development. ED 449 963.

Vykotsky, L. (1978). *Mind and Society*. Cambridge, MA. Harvard University Press.

Wild, L. & Ebbers, L. (2002). Rethinking student retention in community colleges. *Community College Journal of Research and Practice*, 26, (6) 503-519.

Woodward, A., Gosser, D. & Weiner, M. (1993). Problem-solving workshops in general chemistry. *Journal of Chemical Education,* 70, 651-665.

Patrick M. Lloyd is professor of science, Department of Physical Science, Kingsborough Community College, The City University of New York, 2001 Oriental Blvd., Brooklyn, NY 11235. Correspondence concerning this article may be sent to <plloyd@kbcc.cuny.edu>.

Ronald Eckhardt is professor of biology, Department of Biology, Brooklyn College, The City University of New York, Brooklyn, NY 11210.

Acknowledgements: The authors wish to thank Dr. Arthur Zeitlin, Dr. Peter Pilchman, and Dr. Theodore Markus for their contributions to this project.

The work described in this manuscript was made possible by a grant from the National Science Foundation (STEP 0525370).

Anila Asghar, Julie C. Libarkin

Gravity, Magnetism, and "Down": Non-physics College Students' Conceptions of Gravity

This study concentrates on exploring non-science majors' conceptual understanding of gravity and how they use this understanding while solving problems involving gravity.

I. Introduction

This study investigates college students' ideas of gravity in the context of an entry-level geology course in a North American university. The concept of gravity is central to many sciences, and level of understanding of gravity will influence how people apply knowledge from one domain to another. For example, students who believe that gravity only occurs on Earth may have difficulty applying geological principles to concepts in planetary geology. Entry-level college geology instruction assumes an understanding of fundamental physical science concepts such as atoms, friction, gravity, and density. These basic concepts provide an essential foundation for building students' sophisticated understanding of advanced geological concepts. Mass wasting, for example, is a geologic concept that describes the movement of material under the influence of gravity alone. As a consequence, understanding fundamentals of gravity can dramatically impact how students understand and internalize mass wasting concepts taught in a typical physical geology course. We set out to investigate how students enrolled in entry-level geology, most of whom would graduate from college without university-level physics courses, thought about and applied the concept of gravity while solving problems concerning gravity. The repercussions of students' gravity concepts are then considered in the context of non-physics courses, including implications for reform efforts in physics. Data were collected during the second week of an eleven-week term from two courses with an average enrollment of just over 100 students. Based on research literature (Claxton, 1993; diSessa, 1983; Vosniadou, 1994) and one author's experience of teaching this population of students, we hypothesized that students would have fragmented and loosely connected ideas about gravity.

II. Research Context

The investigation of children's and adolescents' ideas about various scientific models and the nature of science is a long-standing and well-developed area of research (Driver, 1985; Driver, Squires, Rushworth, & Wood-Robinson, 1994). Although a few well-known studies have extended this work to include adults, including college students and in-service teachers, there is very little research on the conceptual understanding of gravity on the parts of adults (Gunstone & White, 1981; Hestenes, Wells, & Swackhamer, 1992). In addition, the importance of physical concepts in understanding other sciences suggests that investigation of gravity ideas in related disciplines will have implications for the teaching of physics in both high school and college.

Existing literature on younger students provides a useful backdrop for framing our study of college students enrolled in entry-level geology courses. Gravity-related studies focus on students' ideas of the relationship between gravity/gravitational force and other important concepts in physical science such as height, weight, or velocity. Additionally, some have also examined how children relate orientation ("up" and "down") with reference to the Earth in space to weight and gravity (Driver, et al., 1994; Sneider & Pulos, 1983). Studies have been conducted in K-college settings (Nussbaum, 1976; Trumper & Grosky, 1997) and in a variety of countries (Mali & Howe, 1979; Sneider & Pulos, 1983; Za'Rour, 1975). The majority of

studies conducted in college settings incorporate the concept of gravity into the study of students' force concepts, rather than explicitly studying gravity ideas themselves (Hestenes et al., 1992; Sadanand & Kess, 1990). Kavanaugh and Sneider (2006-2007) provide a thorough review of relevant literature on gravity conceptions; we elaborate on those studies of most importance here.

Conceptual understanding of the source of gravity is an intriguing line of research and is of particular relevance to this paper. *Gravity needs air to act* appears to be a widespread notion among children (Driver et al., 1994; Ruggiero, Cartielli, Dupre, & Vicentini-Missoni, 1985). Similarly, gravity is perceived to be caused by "air pressure" by high school and college students (Hestenes et al., 1992). Elementary and secondary students in several studies also thought that gravity only operated on Earth (Bar, Zinn, Goldmuntz, & Sneider, 1994; Stead & Osborne, 1980). Subjects generally supported this notion through the idea that gravity needs a medium, such as Earth's atmosphere, in which to operate.

The concepts of the geomagnetic field and the Earth's rotation also seem to influence students' ideas of gravity in cross-age models (Stead & Osborne, 1980). A number of investigations focusing on children's and adolescents' understanding of gravity and weight report the prevalence of the idea that gravity somehow only influences heavy things; some 15-year-olds believe that weight is independent of gravity such that objects can have weight without gravity (Driver et al., 1994; Osborne, 1984; Stead & Osborne, 1980).

Some students and even adults thought of "weight as a property of an object," while gravity was perceived as a "property of space," independent of the presence of an object (Ruggiero et al., 1985). In a study of pre-service high school teachers enrolled in a college biology course, Trumper and Gorsky (1997) found that 25-50% of these freshmen through senior college students rarely or never recognized the relationship between weight and gravity. Drawing on Piaget's work (Piaget, 1972), Galili (2001) points out two main schemes involved in children's naïve perceptions of weight: "weight is the pressing force featuring particular objects—the sensed heaviness related to a muscular effort", and "weight is the amount of matter" in an object (Galili, 2001, p. 1085). Several studies report a considerable degree of confusion in children's concepts of weight and gravitation (Galili & Bar, 1997; Galili & Kaplan, 1996; Gunstone & White, 1980; Ruggiero et al., 1985). The inability of secondary and advanced placement students to resolve their misconceptions about weight, gravitational force, and weighing is a consequence of the failure of physics instruction, argues Galili (2001).

Many of the students who held the idea that gravity is a force of attraction had difficulty explaining why gravity exists on Earth, often relying on functional (i.e., gravity's effects) aspects of gravity on Earth, such as gravity "keeps things from floating away," and "holds people on earth."

Conceptions of the reasons that objects fall also provide insight into gravity understanding. Children generally tend to think that objects fall because nothing is holding them up, because the person dropping the object causes the fall, or because things just fall "naturally" (Gunstone & White, 1981; Ruggiero et al., 1985; Selman, Krupa, Stone, & Jacuette, 1982). Adolescents (12 to13 year-olds) were found to have three main non-scientific explanations for falling: (a) an integrated view of gravity and weight where gravity is acting on the weight of a falling object, (b) a clear distinction between gravity and weight, where they both act separately to cause the fall, and (c) an absence of gravity or weight conceptions where objects fall in the absence of support due to "natural motion" (Ruggiero et al., 1985; Selman et al., 1982). A prevalent notion among people of all ages is that heavier objects fall faster because they have a greater acceleration due to gravity (Gunstone & White, 1981; Hestenes et al., 1992; Osborne, 1984; Selman et al., 1982). For example, Gunstone and White (1981) found that approximately 10% of first year physics students hold this belief. In addition, although several studies found that teenagers overall think that gravity decreases with height, a considerable proportion of students thought that gravity actually increases with height above the Earth's surface (Driver, 1985; Driver et al., 1994; Ruggiero et al., 1985).

Research on children's ideas about gravity has also focused on their notions of gravity with respect to spatial orientation. Several studies have reported that children's ideas about "down" with reference to the Earth in space progress through different stages (Baxter, 1989; Driver

et al., 1994; Nussbaum, 1985; Sneider & Ohadi, 1998; Sneider & Pulos, 1981). Children appear to develop an understanding of gravitational down as early as two years of age (Hood, 1995). Children usually start off with an "absolute view of down" that is independent of Earth, eventually developing an "Earth-referenced view of down" at around the age of 14 (Nussbaum, 1985). In a study with 15- and 16-year-olds, Baxter (1989) found that about 80% of students had an earth-referenced view of down based on the Earth's surface rather than the center of the Earth. This Earth-referenced view usually translates into an "up" position for the Earth. Typically, the Northern Hemisphere is "up" relative to the Southern Hemisphere, but we should caution that this view may be dependant upon the Northern Hemisphere populations studied. For example, in two studies in England, Baxter (1989) and Arnold et al. (1995) found that the majority of 7 to 16 year-olds viewed the Northern Hemisphere of the Earth as "up." As a consequence, students believed rain fell away from the Earth, towards space, in the Southern Hemisphere.

A few researchers have investigated pre-teenager (7 to 12 years-old) student conceptions of objects falling through a tunnel in the Earth, as reproduced in this study (Nussbaum 1976, 1979). These investigations are embedded in larger studies of student understanding of Earth's shape and gravity, with a primary purpose of understanding the general gravity orientation that students hold. Three conceptual orientations are considered in existing studies: (a) objects falling towards the center of the Earth, (b) objects falling towards the surface of the Earth, and (c) objects falling down in space. Generally, these studies consider a drawing depicting an object falling and stopping immediately at the Earth's center as "correct", rather than the more accurate model of an object oscillating about the Earth's center. Using a series of activities prompting elementary-aged students to depict the path taken by a rock, Nussbaum (1976) found that young students often have perspectives of "down" that are reference frame specific. Students drew rocks (a) falling towards the bottom of the page, (b) falling towards the Earth's surface but not the center, (c) falling away from the Earth, and (d) falling towards the Earth's center. Nussbaum (1979) extended this work by providing students with four drawings and asking them to choose the one that they thought best represented what would happen to an object falling through the tunnel. These results were similar to those of earlier work, and revealed that the gravity orientation of most young students is not in reference to the Earth's center. The studies tend to explain students' responses to these problems, but do not attempt to probe the ways in which students understand the problems. Students' mental images or perceptions of the tunnel are not clear from existing work; for example, do students think that the tunnel has air or water in it, or do they consider variables such as the Coriolis effect?

Through this and other future studies we hope to engage in a conversation about these issues across the disciplinary boundaries that may exist between physicists, other scientists, and science educators.

In related work, Sneider and Ohadi (1998) investigated elementary and middle school students' conceptions of shape and gravity in the context of evaluating the effectiveness of an instructional intervention. This study found that, prior to instruction, only 7-30% of 4th through 8th graders thought that objects fall towards the Earth's center. After instruction, only 60% of 8th grade students acquired this concept. This suggests that some alternative ideas about the orientation of gravity are entrenched and may be prevalent even in older students.

Although there is ample existing literature on children's and adolescents' (K-12 students) ideas about gravity, there is a dearth of literature on college students' and pre-serivce teachers' understanding of gravity. This study specifically concentrates on exploring non-science majors' (typical undergraduate students including pre-service teachers and child development majors) conceptual understanding of gravity and how they use this understanding while solving problems involving gravity. More specifically, we examine the ways in which student ideas resonate or conflict with the scientifically accepted idea of gravity. Additionally, we seek to closely examine alternative frameworks about the nature of gravity through a qualitative analysis of student written and pictorial responses. In particular, we wanted to answer two questions: (1) How do non-science major college students enrolled in a science course other than physics understand gravity? (2) In what ways do they apply their understanding of gravity while solving problems related to gravity? Additionally, we looked at the relevant literature to compare the participants' understanding of gravity to models of gravity held by

K-12 students. For the purpose of this study, we defined gravity very simply as the force of attraction between two or more objects that have mass.

III. Methods

The narrowly constrained research focus of this study allowed for implementation of methodologies used previously by other researchers. In particular, we adopted gravity problems that Nussbaum (1976, 1979), had originally used with elementary students for use with our non-science major college population. We specifically targeted students enrolled in geology courses in which an understanding of physics concepts is generally assumed.

A. Participants

The study population (n=216) consisted of students enrolled in two entry-level university geology courses taught by one of the authors during winter 2004 (n=96) and fall 2005 (n=104) at a mid-western North American university. The two courses participating in this study were almost evenly divided between men and women, with a total of 117 men and 99 women enrolled in the courses overall (48% female). Students had mixed levels of college experience. 49.5% of the students were in their first year of college. 2nd, 3rd, and 4th/5th year students made up 28%, 13%, and 9% of the study population, respectively. At the start of the term, 28% of the students were education or early childhood majors, including one Earth Science Education major, 20% were undecided, and six students had declared Geosciences as their major. Remaining students were enrolled in diverse majors, including finance, sociology, English, plant biology, forensic chemistry, mechanical engineering, and art. Of these majors, 17% were declared as math, science (including geo-science), computer science, or engineering majors (Table 1).

Table I: Course and participant demographics.

	Female	Male	1st-year students	Education or Child Development majors	Undecided	Geology majors
Winter 2004	60	46	57	36	20	0
Fall 2005	44	66	50	25	22	6
TOTAL	104	112	107	61	42	6
Participants	104	104	—	—	—	—
% participation	100%	92.8%	—	—	—	—

The response rate to the questionnaire from both courses was 96% (n=208). The institution where this study was conducted is situated in the rural mid-west, although the majority of students are from suburban homes near small cities. The collection of gravity conceptions data from students was conducted with a dual purpose: as part of instructional pre-assessment and as conceptions research. All students enrolled in the courses participated in the pre-assessment. Attendance rates for F2005 and W2004 classes were 100% and 93%, respectively. 68% of students enrolled in these courses reported previous high school Earth Science or Physics courses, suggesting a significant level of exposure to gravity concepts in high school, although instruction on gravity occurred in this course only after these data were collected.

B. Data Collection

The questionnaire utilized in this study included an open-ended question and two gravity tasks. The open-ended question asked students to explain what gravity is and the cause of gravity on Earth (What is gravity? Why is there gravity on Earth?). The gravity tasks had previously been investigated with elementary-aged children (Nussbaum, 1976, 1979). The gravity tasks asked students to pretend that a tunnel was dug all of the way through the Earth from the North to the South Pole, imagine that a person standing at the surface dropped a rock, and draw a line from the person's hand showing the path taken by the rock. Students were also asked to explain their response. In a second task, the frame-of-reference of the first task was modified to show a tunnel oriented along the equatorial plane (Arnold, Sarge & Worrall, 1995) in order to allow for investigation of the reference dependence of gravity concepts. This allowed differentiation between those students who believed objects fall towards the center of the Earth and those who believe objects fall towards the bottom of the page.

The questionnaires were distributed and collected before instruction and had no impact on students' grades. In order to provide for anonymity,

questionnaires were numbered before analysis. 96% of the 216 students enrolled in the courses completed some portion of the questionnaire. Specifically, 208 students completed the drawing tasks, and 197 completed the open-ended gravity question. These data provide a unique look into the gravity conceptions of typical college students, as well as a large subset of pre-service teachers and child development majors.

This study did not probe students' mental picture of the north-south and east-west tunnel, although, based on the written responses, we have no indication that students are considering phenomena such as the Coriolis effect. We feel that future studies would benefit from probing during interviews to elicit a more accurate and richer picture of students' ideas.

C. Data Analysis

Each of the researchers analyzed one dataset completely, and conducted repeat analysis on the other dataset to establish inter-rater reliability as discussed below. We analyzed the student responses and the drawings. Questionnaire responses were analyzed via thematic content analysis (Patton, 1990), wherein themes are allowed to emerge naturally from the data. The data were divided into two sets for the purpose of analysis: (a) responses to the open-ended gravity question and (b) drawings related to the gravity tasks. Students' qualitative responses to the open-ended question and explanations of their gravity task drawings were tabulated electronically. The responses were coded thematically to capture important ideas and misconceptions expressed by the participants in relation to gravity. Codes were grouped into broader categories and general themes as shown in Table II. Responses to each of the two gravity tasks were also analyzed and coded thematically, with dominant categories of responses emerging from the data.

Table II: Common alternative conceptions of gravity identified in the study population.

Alternative conceptions	Exemplar student responses (Gravity is…)
Gravity as an outside force	• force acting on the objects of the universe in varying degrees
Gravity associated with the atmosphere	• nature's force pulling down because of the atmosphere
Gravity associated with celestial objects and/or earth's spin	• a force created by objects in space, there is gravity on earth because it is spinning in space • There is gravity on earth because of rotation of the earth • the pull or push of mass towards the earth. Because of Earth's mass and constant rotation and the pull of the moon
Gravity associated with the sun	• a force that pulls everything down. There's gravity on Earth because of its relationship to the sun
Gravity as the pull or attraction of Earth's core	• the pull the earth's core has on everything • Force of attraction from the center of the planet that holds things together. If there was no gravity, things would not be held together and everything would just float into space and not stay on Earth
Gravity associated with pressure	• the pressure that is forced onto an object at 9.8m/s^2
Gravity associated with size and earth's position	• a force that pulls on all objects. Depending on the size of the planet, gravity's pull is stronger. Caused by [earth's] position in solar system and closeness to sun
Gravity associated with magnetism	• a force that holds everything on the earth, gravity is here because the earth spins and revolves and maybe because its magnetic • a magnetic force caused by the star or the sun. • a force within the earth caused by metallic substances rotating in the molten core, this also affects the magnetic field… • There is gravity on Earth because of the magnetic field of the solar system. Gravity is related to mass
Functional definition of gravity on Earth	• the force that holds everything in place on Earth to keep everyone in place and everything • a force that pulls on objects. Gravity is here to keep everything in its place on earth otherwise we would just float away with no control

1. Validity and reliability

The gravity tasks included in this questionnaire were adopted from previous studies[23] that have already undergone validation for elementary-aged students. We view this prior use and validation as a significant measure of the usability of these tasks for eliciting gravity concepts. The application of materials developed for K-12 to the college population required further validation. Both the drawing tasks and the open-ended gravity questions created for this study were first used in a pilot study. This piloting occurred in a similar entry-level geology course taught before the courses included in this study. Responses to the pilot questionnaire indicated that students were able to understand and respond to the tasks and questions. In addition, college students did not feel that these tasks were too simplistic for a college level course.

Additional aspects of qualitative validity were addressed during this study. Students were also consulted through in-class discussions to obtain their feedback on the analytic categories to address the credibility of the interpretations made by the researchers (Patton, 1990). Students in both courses agreed that their personal models fell within one of the described models. This process helped in ascertaining the level to which the study participants agreed with the research findings. For the purpose of creating inter-rater reliability, each researcher also coded 20 responses from the alternate dataset. The scores and codes were compared and will be discussed later in this article. The inter-rater agreement was 100% for drawings, and was 80% to 100% for open-ended responses, based on discussion of thematic codes. The context of this study and the assessment tool may have influenced student thinking; however, we noticed similar responses on both the written explanation and the drawing task.

IV. Findings And Discussion

In this section, we discuss the salient conceptions emerging from students' responses to the gravity problems. More specifically, student notions of gravity and causes are analyzed, as well as responses and explanations for the gravity tasks.

A. Notions of Gravity

The open-ended questions were used to elicit students' ideas about gravity. Student responses (n=197) revealed a diverse and complex array of notions about both gravity and the relationship between gravity and the Earth (Table II; Figure 1). It is particularly important to note that the majority of the students' responses exhibited a combination of misconceptions about gravity. We did not notice any truly coherent framework emerging from the responses, except for perhaps the scientifically accepted model held by a very small proportion of students. Although we have identified some salient themes arising out of students' responses for the purpose of this analysis, we are not suggesting that students hold only these distinct ideas. In fact, the data suggest that a multitude of alternative ideas co-existed in participants' minds, albeit mostly reflecting incoherent frameworks.

Students tend to reason from familiar concepts, and when these concepts are non-scientific, conclusions can be far removed from the concepts we think we are teaching.

1. Gravity as a force of attraction

Students' responses suggested that, surprisingly, only 21% of the participants had the correct scientific idea that gravity is a force of attraction. The specific idea that the force of attraction exists between masses was ultimately not coded as very few students incorporated all three concepts of attraction, mass, and force in their explanations. Many of the students who held the idea that gravity is a force of attraction had difficulty explaining why gravity exists on Earth, often relying on functional (i.e., gravity's effects) aspects of gravity on Earth, such as gravity "keeps things from floating away," and "holds people on earth." Furthermore, merely 12% of students held the correct conception that gravity is a force without simultaneously holding common alternative conceptions. The 9% of the study population with dual scientifically accepted/non-scientific models held misconceptions that were mostly in relation to why gravity exists on Earth. A number of students thought that Earth's "rotation," "spin," "magnetism," and "atmosphere" cause a force of attraction between the Earth and other objects. In fact, one student mentioned all of these concepts in her definition of gravity, explaining that gravity is a force of attraction and that "… Gravity is a force due to spinning of the earth around the sun, magnetism, and atmosphere". The following selected excerpts illuminate students' reasoning for various misconceptions:

"Gravity exists on the earth, because of Earth's mass and constant rotation and the pull of the moon."

"Gravity is on earth to keep things on earth."

"There is gravity on earth because of the earth's core."

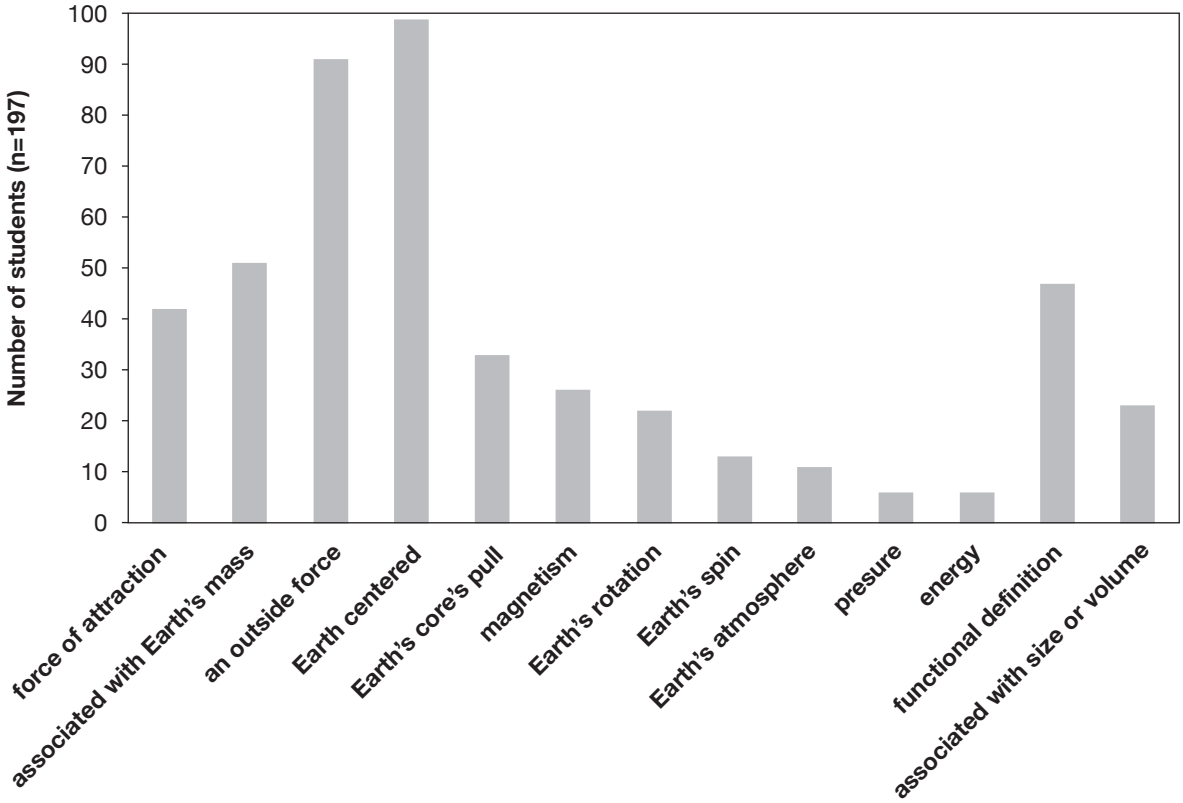
2. Gravity as an outside force

Approximately 46% of the participants seemed to view gravity as an outside force acting on an object, which might (a) exist independently of the object and (b) originate from something other than the Earth plus the attracted object. For example, 93% of those who perceived gravity as an outside force associated it with the Earth specifically. As observed in other categories, responses reflected multiple alternative conceptions about gravity. Not only did students think of gravity as a force acting on the Earth, they also suggested that gravity exists on Earth because of "magnetism," "earth's rotation," "earth's spin," "atmospheric pressure", and Earth's "relative position to the sun". About 29% of students in this category also mentioned the function of gravity on Earth (i.e., gravity prevents objects from "floating away"), either citing it as the sole reason for the existence of gravity on earth or coupling it with other phenomena, such as magnetism and Earth's rotation. A few students also associated gravity with the movement of objects. A large number of those who thought of gravity as an outside force acting on objects also invoked several other concepts that may or may not be connected to one another. The following responses reflect the diverse ideas these students carried in relation to gravity:

"Gravity is a force, acting on earth keeping everything on the ground. There is gravity because of the position in the solar system as well as the rotation of Earth."

"Gravity pulls on everything keeping it in place. There is gravity on Earth because of the way the Earth rotates."

Figure 1: College student conceptions of gravity and its causes. Alternative conceptions as illustrated in Table II and additional ideas identified in the study population. Notice that over 90% of the students in this study had an earth-centric perspective of gravity, and believed that gravity was a force acting on the Earth, rather than something that was an inherent characteristic of Earth (and all objects with mass). A wide array of other alternative conceptions were observed; see text for details.



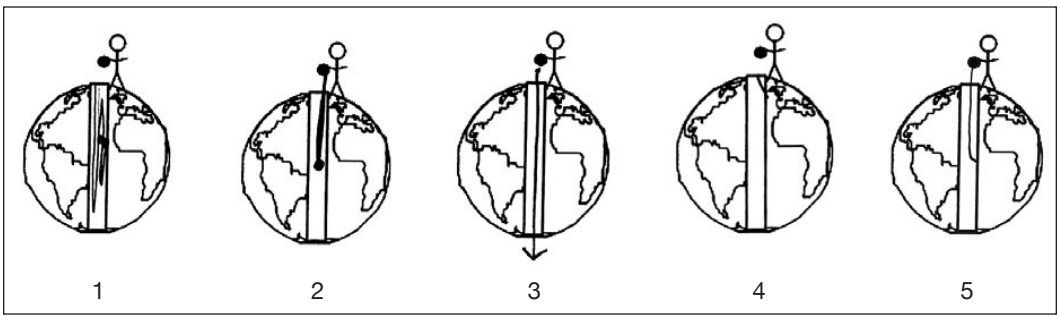
"Gravity is the force put upon earth due to earth's rotation around its axis and the sun."

"Gravity is the force pushing things down on the earth. Attraction of one object to another. Energy that pulls. Large objects have stronger [gravity]. With out [sic] [gravity] things would float."

"It's a magnetic like force that pulls everything towards the earth. If we didn't have it every thing [sic] would float off the earth that wasn't held down."

Although these five participants' responses seemed to imply that gravity is an external force "acting on earth" that "pulls" or "pushes" all objects on the Earth towards the "center of the earth" in order to "keep" objects "in place", they tended to give divergent explanations for why gravity exists on Earth. Two participants said that gravity is caused by the "rotation" of the Earth, and one of them also cited the relative "position" of the Earth in the solar system. One student said that Earth's spin and the "sun" caused gravity on the Earth. Two students mentioned only the function of gravity (i.e., without gravity things would "float away") in their explanations to account for gravity's existence on the Earth. One of these students also mentioned gravity as the force of "attraction" between objects while simultaneously viewing it as an external force "pushing" objects down on the Earth. Additionally, not only did this student define gravity as a "force," she also described it as "energy that pulls." She was applying the concepts of force and energy to gravity in an incoherent way, which suggests a faulty understanding of these concepts as well. Participants' responses suggest that they may be using several technical terms, such as "energy", "magnetic force", or "spin", to explain their ideas of gravity without a clear understanding of what these terms mean.

Figure 2: Exemplars of categories of student drawings for a N-S oriented tunnel. Category 1 is the correct conception, depicting the rock oscillating about the Earth's center. Category 2 is the most commonly observed conception, where the rock follows a direct path and stops at the Earth's center. Category 3 represents the concept that the rock will travel to the opposite side of the Earth. Three variations of category 3 were observed: the rock falls out the opposite side of the Earth, shown here, the rock will stop at the opposite side of the Earth, and the rock will fall out the opposite side of the Earth, curve and hit the Earth's surface. Category 4 and Category 5 were related but distinct concepts. Category 4 is the idea that the rock will curve and come to rest at the side of the tunnel very early in its journey down the tunnel. Category 5 blends Categories 2 and 4, yielding the notion that the rock will fall to the center of the Earth and curve into the side of the tunnel.



3. Earth-based notion of gravity

About 50% of the students in this study seemed to think of gravity only in relation to the Earth, and most held other alternative ideas along with this idea. As examples, students described gravity as "a force that the earth is exerting on us", "a pull the earth's core has on everything", "a force that earth's mass creates because of its size", or a "magnetic pull caused by the earth's core". A number of participants also correctly connected gravity on Earth to Earth's mass in their responses. As one student put it, gravity is a "pull on an object towards the center of the earth by the earth … the mass of earth in the solar system creates the pull".

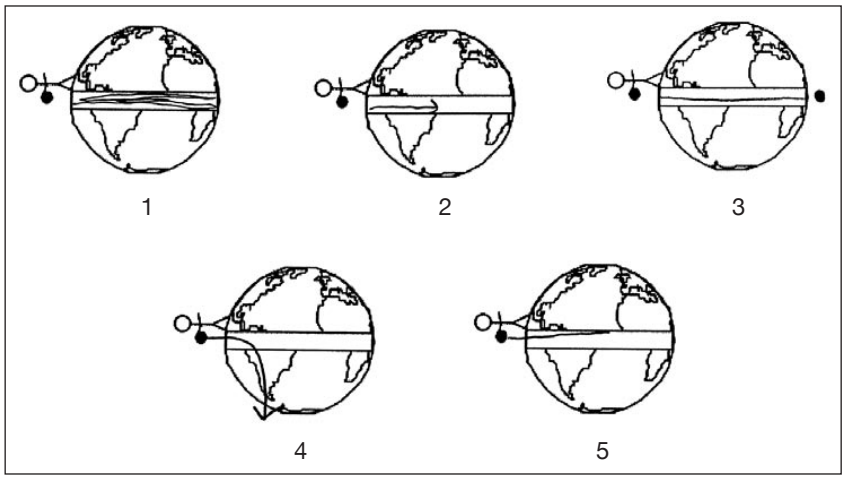
Again, several participants mentioned magnetism, Earth's rotation, Earth's spin, or a functional view of gravity to explain why gravity exists on Earth. A few mentioned the Sun's and "moon's pull" as reasons for Earth's gravity. Surprisingly, some students thought of gravity as "energy" "pulling on" objects. According to one student, gravity is "energy that pulls you." Interestingly, some used the concept of Earth's "size" instead of mass while explaining their understanding of Earth's gravitation.

Student application of their understanding of gravity to the two simple gravity tasks also suggests a limited understanding of gravity. Overall, 73% of students drew the same pathway for both the N-S and E-W tunnels, and the likelihood of holding the same model regardless of tunnel orientation depended upon

the student's overall understanding of gravity. Conceptions were evenly distributed across the two studied course populations.

Five categories of rock paths dominated the student responses for both N-S and E-W tunnel orientations, and two additional categories of response are worth discussion (Figure 2; Figure 3; Table III). Category 1 drawings depicted a rock oscillating about the Earth's center, which was the correct idea, and all students (10%, n=21) who held this idea for the N-S tunnel also held it for the E-W tunnel. One additional student held a Category 1 idea for the E-W tunnel and a Category 2 idea for the N-S tunnel. Category 2 drawings showed a rock path that stopped abruptly at the center, and these were the drawings most commonly observed. Overall, 49% (n=98) of students held this idea for both tunnel orientations, six held it for the E-W tunnel only, and thirteen students expressed this idea for the N-S tunnel only. These students were mixed between Categories 3, 4, and 5 in their choice for the other orientation or believed the rock in the E-W tunnel would simply fall off the page (Figure 4). Category 3 drawings depicted a rock that fell through the Earth. Once on the other side, the rock fell out of the tunnel, fell out and curved back down to the Earth's surface, or simply stopped moving. While 45 students held this model for the N-S tunnel, only 20 drew similar paths for both tunnels. Category 4 and 5 were related, and both showed the rock curving and coming to rest at the side of the tunnel. Category 4 paths ended near the tunnel entrance, while Category 5 paths ended at the Earth's center. Nearly twice as many students believed the rock would curve towards the tunnel surface for the E-W oriented tunnel (n=31) than the N-S tunnel (n=18).

Figure 3: Exemplars of categories of student drawings for the E-W oriented tunnel. Categories as in Figure 2.



Category 6 and 7 ideas (Figure 4) were held by fewer students than other models and were observed predominantly for specific tunnel orientations. Four students indicated that the rock in the N-S tunnel problem would fall away from the Earth, "slide" around the Earth, or orbit the Earth. An additional student indicated that, while he believed most strongly that the rock would fall to the opposite side of the tunnel and stop (Category 3), the rock moving around the Earth instead of into the E-W tunnel was a possibility. Another student preferred the model of the rock stopping at the center of the N-S tunnel (Category 2), but suggested that the rock might also "fall straight down due to gravity". This explanation was accompanied by a crossed-out path that showed the rock moving perpendicular to the N-S tunnel (that is, off the page). Finally, twelve students believed that the rock would fall off of the page rather than move into the E-W tunnel. Of these, two also held the conception that the rock would move away from or around the Earth for the N-S tunnel orientation (Category 6).

Student explanations of their drawings did not always indicate deep understanding or firmly held beliefs, but they provided interesting insight into student thought processes. Although Category 1 represented a correct concept, students were generally unable to provide a detailed explanation of this phenomenon or demonstrate alternative ideas. For example, our Category 1 student in Figure 2 explained his drawing by stating "The center pulls it there by 9.8 m/s and it rubberbands back + forth till [sic] it gets to the center". This student depicted an identical path for the E-W tunnel, stating simply, "Earth is round". Another student explained this oscillating phenomenon in the N-S tunnel with an alternative conception about the cause of the attraction between the rock and the Earth, stating "Gravity and earth's rotation will balance out".

The dominant conception that the rock will follow a straight path and come to rest at the Earth's center was often explained by a tug-of-war between the two ends of the tunnel ("The rock will not fall the entire way through the center of the earth b/c the gravity will pull the rock in the center of the 2 ends") or by the perceived attractive nature of the core ("core pulls

things towards it"). The final motion of the rock at the end of the tunnel, whether continuing to move into space, falling to the Earth's surface, or simply stopping, often changed as the tunnel changed orientation. A few students used their explanation for the rock's path in the N-S tunnel to justify a similar depiction for the E-W tunnel ("He is still standing next to a hole and when he drops it it will still fall straight down through the tunnel"). One student who believed the rock would fall straight through the N-S tunnel and out the other side ("The rock will fall until gravity is unable to act upon it without another for [sic] acting up on it) also believed the rock would fall off of the page rather than into the E-W tunnel ("The rock will fall and have gravity enforced on it until there is no more gravity"). Eight of the twelve students with the "off of the page" conception believed the rock would fall all of the way through the N-S tunnel to the other side. The remaining four students were evenly divided between a rock that stopped at the center of the N-S tunnel and a rock that curved away from the Earth or began to orbit the Earth (Figure 4; "Because of the earths rotation the rock would slide away from the tunnel"). Finally, students with Categories 4 and 5 models believed that gravity pulls objects down ("The rock would go into the Earth but then move downwards because of the gravity pulling it down") or towards a surface ("Gravity would eventually draw it to some sort of surface"). One student emphatically described her Category 4 conception as shown in Figure 2, "The rock will fall quickly to the Earth. It will <u>not</u> go through the tunnel". Interestingly, for most students the surface in the tunnel to which the rock falls could be orientated vertically or horizontally relative to the page. It appears that these students had an earth-referenced conception of up/down relative to the surface of the Earth.

A few students provided explanations of the rock's path that fell outside of the categories described here. One student believed that the rock would fall into the tunnel and come right back to the surface for both tunnel orientations. This student did not have an explanation for this phenomenon and simply stated that he "saw it in a magazine". Four students believed the rock would not move at all in the E-W tunnel, with one of these students holding the identical belief for the N-S tunnel. The student with the "no motion" conception for both tunnels firmly stated that "the rock won't move" without further explanation. The remaining three students with no motion in the E-W tunnel all held the conception that the rock would follow a direct path in the N-S tunnel and stop at the Earth's center. These students explained their 'no motion' model in various ways. One student thought the rock would "fall to ground" but not into the tunnel. Another student similarly explained that the rock would "drop and stay in that area", and a third student indicated that the rock "would just stop".

Table III: Seven categories of student drawings.

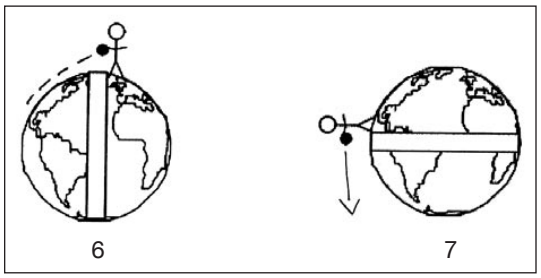
Drawing Category	N-S Tunnel # of students (% of n=201)	E-W Tunnel # of students (% of n=201)
Category 1 drawings depicted a rock oscillating about the Earth's center, and this represents the correct idea.	21 (10%)	22 (11%)
Category 2 drawings depicted a rock that stopped abruptly at the center.	111 (55%)	104 (52%)
Category 3 drawings depicted a rock that fell through the tunnel to the opposite side of the Earth.	45 (22%)	20 (10%)
Category 4 drawings depicted a rock that fell into the tunnel (not to the center) and curved to the tunnel's surface.	5 (2%)	15 (7%)
Category 5 drawings depicted a rock that fell to the Earth's center and curved to the tunnel's surface.	13 (6%)	16 (8%)
Category 6 drawings depicted a rock falling away from, or curving around, the Earth.	3 (1%)	0
Category 7 drawings depicted a rock falling off the page.	0	12 (6%)
Total number of responses in these categories	196	189

C. Undergraduate versus K-12 Students' Conceptions of Gravity.

This study was an attempt to explicitly link gravity conceptions and problem-solving approaches of K-12 students with those of non-physics college students. Surprisingly, college students' conceptions of gravity seem remarkably similar to conceptions identified in younger populations, particularly elementary-aged students. In particular and as described earlier, the research literature abounds with non-scientific models of gravity held by children. For instance, children tend to think that gravity either needs air to act or is caused by air pressure. Another prevalent misconception is that gravity operates on Earth only. Many K-12 students tend to associate gravity with magnetism or Earth's rotation. An "Earth-referenced view of down" based on the Earth's surface rather than the center of the Earth is also prevalent among children. While solving the gravity task in a similar study, elementary and middle schools students also drew rocks (a) falling towards the bottom of the page, (b) falling towards the Earth's surface but not the center, (c) falling away from the Earth, and (d) falling towards the Earth's center (Nussbaum & Novak, 1976; Nussbaum, 1979). Alarmingly, college student misconceptions are not dramatically different from those of K-12 students, and a significant number of students in our study population held ideas similar to those of elementary students. The notions that gravity exists because of magnetism, air pressure, the special nature of the core, or the Earth's rotation are all well-documented in pre-college populations and were also observed in abundance here.

Overall, the solutions and explanations that college students provided for the gravity tasks were generally more advanced than observations made of elementary students in the 1970s (Nussbaum, 1979). However, over 5% of the studied college population held the idea that gravity worked down relative to a piece of paper, and 21% believed a ball would simply fall out of a N-S oriented tunnel and into space. Almost 50% of the college students mentioned in their responses that a dropped object will immediately stop at the Earth's center; this was classified as a "correct" model by earlier researchers investigating young children. Finally, very few students were able to provide coherent explanations for object behavior, suggesting that college students hold incoherent ideas about gravity and related phenomena.

Figure 4: Other common rock paths produced by the study population were dependent upon tunnel orientation. Category 6: Four students depicted the rock falling away from the Earth or curving around the Earth. This was only observed as a primary model for the N-S oriented tunnel, although one student suggested this as an alternative model for the E-W tunnel. Category 7: Twelve students indicated that the rock would fall off of the page for the E-W oriented tunnel only.



Despite more than a decade of diverse national efforts to modify and improve physics instruction in the K-college classroom, we observed little difference between current college student models and those of young children documented almost three decades ago. For example, only 10% of our college population held the simplest scientifically accepted model of gravity without also simultaneously holding alternative ideas, suggesting little change in gravity understanding from K-12 to college. The alternative ideas observed here were quite similar to those of elementary students, and a high percentage of both groups of students equated gravity with magnetic or rotational forces (Table I; Figure 1). Ideas of gravitation that have long been considered to be exclusive to young students, such as the "off-the-page" model classified here as Category 7 (Figure Z) and documented by Nussbaum (1979), were also observed in our adult population. These findings suggest that gravity models are either resistant to change, are prone to reversion from scientific models to original alternative conceptions, or are not being explicitly targeted by instruction.

Student responses to the gravity problems also revealed a host of alternative ideas about other fundamental physical and geological concepts. The connection of gravity with magnetism, Earth's rotation, atmospheric pressure, and the core of the Earth suggests a limited understanding of these phenomena. Other studies also suggest that many children and adults relate air, atmosphere, magnetism and Earth's spin to gravity (Ruggiero et al., 1985; Stead & Osborne, 1980). Additionally, half of the students in this study believed that a falling rock would reach the center of the Earth and stop, and many students explained this by saying that there would be "no force" at the center. While the students are correct

that there would be no net force at the center of the earth, they are incorrect in their depiction of the rock stopping at the center. This indicates that these college students may be confused in their usage of the concepts of force, velocity, and acceleration. It appears that students need to develop a clear understanding of force and motion to understand gravity as a force of attraction that is mediated by a field. These data bring to mind a number of interesting questions about the relationships between physics reform efforts, pre-service preparation, in-service continuing professional development, and instruction in K-12 classrooms. Through this and other future studies we hope to engage in a conversation about these issues across the disciplinary boundaries that may exist between physicists, other scientists, and science educators. What would it take to help students and teachers to develop coherent models of foundational ideas in physics? How do we supplement or follow up on "good teaching" practices so that students retain content and develop a coherent and lasting understanding of scientifically accepted models?

V. Conclusions and Implications

We found that entry-level college students enrolled in geology courses were unable to provide any coherent scientific explanation of gravitationally related phenomena. Neither weak explanations, such as the simple regurgitation of the gravity definition, nor strong explanations, including those that connect meaningfully to other scientific ideas, were present in our population. Students' own ideas about gravity reflect a lack of understanding about many fundamental ideas in science, including pressure, gravity, magnetism, and Earth's rotation (Figure 1). Additionally, non-physics college student conceptions of gravity are remarkably similar to ideas held by elementary-aged students reported nearly three decades ago. This suggests that physics instruction in pre-college grades may need to explicitly address gravitation ideas held by students, rather than more common approaches that focus on the effects of gravity. The impact of alternative ideas on reasoning in other domains, such as Earth and Space science, cannot be understated. For example, the Earth-centric models of gravity held by many of the students in our study resulted in some who believed that gravity only existed on Earth. Other students (in related work with this same population of students), felt that gravity did not exist on Mars, because Mars does not have a magnetic field. The likelihood that students will confuse these foundational ideas with other conceptions suggests that curriculum developers may need to identify and address students' prevalent misconceptions in physical science by including multiple experiences that help them to transform and develop their initial ideas into coherent explanations of these basic scientific ideas.

Students tend to reason from familiar concepts, and when these concepts are non-scientific, conclusions can be far removed from the concepts we think we are teaching. It then becomes paramount that we explicitly target the most basic ideas in our instruction, rather than assuming, as certainly many college faculty do, that students have "already had that material" in an earlier class. The findings presented here suggest that college faculty in all sciences, not just physics, may want to identify, discuss, and clarify students' confusions related to gravitation, weight and other related concepts as a core element of effective instruction. For geoscientists, misunderstanding the fundamentals of gravity can have serious implications for teaching any number of larger ideas, including mass wasting, plate tectonics, and planetary geology. Furthermore, many geology majors are required to take geophysics classes, since geophysics is a fundamental tool used in oil exploration, mining, and other geologic industries. Some of the students in our study were enrolled as geology majors, and at the very least, these majors should understand gravity in order to be able to reason about gravity anomalies and planetary systems. This study suggests that for faculty in interdisciplinary fields such as geology, it may be a good idea to probe about fundamental ideas before engaging in a discussion of concepts that require an a priori understanding.

One of the main limitations of this study is that it was based on a relatively simple tool to elicit college students' ideas about gravitation. We employed this tool to replicate studies originally conducted with children with the college level population. We propose that future research and effort in teaching gravity concepts may need to focus on 1) Interview-based probing of high school and college students' understanding of the concepts of gravitational force, weight, the inertial mass, and the gravitational mass (Gönen, 2008) to attain a deeper knowledge of their mental images and conceptions of these ideas; 2) Conducting pre-post and experimental studies to examine the effect of instruction on students' prior knowledge of gravitation; and 3) Investigating how teachers translate

concepts learned via reform efforts to K-12 classrooms. K-12 teachers are exposed to reform-based methods, but are not always able to translate these experiences to classroom instruction. Certainly, we know that reform-based courses are effective at engendering conceptual change (Shaffer & McDermott, 2005), but it is less certain if these approaches are being applied effectively (Le et al., 2006). We also wonder if reform-based instruction targets the most common misconceptions about the natural world (Vosniadou & Brewer, 1992). Finally, it might be useful to consider these fundamental concepts from the theoretical perspective of threshold concepts (Meyer & Land, 2005). Threshold concepts are ideas that can act as barriers to a curriculum, wherein understanding a threshold concept may provide a gateway to deep understanding of related concepts. We as faculty may need to seriously reconsider the assumptions we make about students, in terms of both what students are bringing to and taking from the classroom. Ultimately, we as a community need to decide which concepts are absolutely necessary for deeper learning in a domain, and teach accordingly.

References

Arnold, P., Sarge, A., & Worrall, L. (1995). Children's knowledge of the earth's shape and its gravitational field. *International Journal of Science Education, 17*(5), 635-641.

Bar, V., Zinn, B., Goldmuntz, R., & Sneider, C. (1994). Children's concepts about weight and free fall. *Science Education, 78*(2), 149-169.

Baxter, J. (1989). Children's understanding of familiar astronomical events. *International Journal of Science Education, 11*(special issue), 502-513.

Claxton, G. (1993). Minitheories: a preliminary model for learning science. In P. J. Black & A. M. Lucas (Eds.), *Children's Informal Ideas in Science* (pp. 45-61). London: Routledge.

diSessa, A. A. (1983). Toward an epistemology of physics. *Cognition and Instruction, 10*(2 & 3), 105-225.

Driver, R. (1985). *Children's Ideas in Science*. Milton Keynes, UK: Open University Press.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas*. London: Rutledge.

Galili, I. (2001). Weight versus gravitational force: Historical and educational perspectives. *International Journal of Science Education, 23*(10), 1073-1093.

Galili, I., & Bar, V. (1997). Children's operational knowledge about weight. *International Journal of Science Education, 19*(3), 317-340.

Galili, I., & Kaplan, D. (1996). Students' operation with the concept of weight. *Science Education, 80*(4), 457-487.

Gönen, S. (2008). A study on student teachers' misconceptions and scientifically acceptable conceptions about mass and gravity. *Journal of Science Education and Technology, 17*, 70-81.

Gunstone, R., & White, R. T. (1981). Understanding of gravity. *Science Education, 65*(3), 291-299.

Gunstone, R. F., & White, R. T. (1980). A matter of gravity. *Research in Science Education, 10*, 35-44.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141-158.

Hood, B. M. (1995). Gravity rules for 2-to 4-year olds? *Cognitive Development, 10*, 577-598.

Kavanagh, C., & Sneider, C. (2006-2007). Learning about Gravity I. Free Fall: A Guide for Teachers and Curriculum Developers. *Astronomy Education Review, 5*(2), 21-52.

Le, V., Stecher, B. M., Lockwood, J. R., Hamilton, L. S., Robyn, A., Williams, V. L., et al. (2006). *Improving mathematics and science education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement*. Arlington: RAND Corporation. Document Number)

Mali, G. B., & Howe, A. (1979). Development of earth and gravity concepts among Nepali children. *Science Education, 63*(5), 685-691.

Meyer, J. H. F., & Land, R. (2005). Threshold concepts and troublesome knowledge (2): epistemological considerations and a conceptual framework for teaching and learning. *Higher Education, 49*, 373-388.

Nussbaum, J. (1979). Children's conceptions of the earth as a cosmic body: a cross age study. *Science Education, 63*(1), 83-93.

Nussbaum, J. (1985). The Earth as a cosmic body. In R. Driver, E. Guesne & A. Tiberghien (Eds.), *Children's ideas in science*. Milton Keynes: Open University Press.

Nussbaum, J. (1976). An assessment of children's concepts of the earth utilizing structured interviews. *Science Education, 60*(4), 535-550.

Osborne, R. (1984). Children's dynamics. *The Physics Teacher, 22*(8), 504-508.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2 ed.). Newbury Park: Sage

Pfundt, H., & Duit, R. (1994). Bibliography: Students' Alternative Frameworks and Science Education. Kiel, Germany: University of Kiel Institute for Science Education.

Piaget, J. (1972). *The Child's Conception of Physical Causality*. Totowa, New Jersey: Littlefield, Adams.

Ruggiero, S., Cartielli, A., Dupre, F., & Vicentini-Missoni, M. (1985). Weight, gravity, and air pressure: Mental representations by Italian middle school pupils. *European Journal of Science Education, 7*, 181-194.

Sadanand, N., & Kess, J. (1990). Concepts in force and motion. *The Physics Teacher* 530-533.

Selman, R., Krupa, M., Stone, C., & Jacuette, D. (1982). Concrete operational thought and the emergence of the concept of unseen force in children's theories of electromagnetism and gravity. *Science Education, 66*, 181-194.

Shaffer, P. S., & McDermott, L. C. (2005). A research-based approach to improving student understanding of kinematical concepts. *American Journal of Physics, 73*(10), 921-931.

Sneider, C., & Ohadi, M. (1998). Unraveling students' misconceptions about the earth's shape and gravity. *Science Education, 82*, 265-284.

Sneider, C., & Pulos, S. (1981). Children's cosmographies: understanding the earth's shape and gravity. *Science Education, 67*(2), 205-221.

Sneider, C., & Pulos, S. (1983). Children's cosmographies: Understanding the Earth's shape and gravity. *Science Education, 67*, 205-222.

Stead, K., & Osborne, R. (1980). A working paper of the learning in science project. Hamilton, New Zealand: University of Waikato, Science Education Research Unit.

Trumper, R., & Grosky, P. (1997). A survey of biology students' conceptions of force in pre-service training for high school teachers. *Research in science and technological education, 15*(2), 133-147.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4*, 45-69.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*, 535-585.

Za'Rour, G. I. (1975). Science misconceptions among certain groups of students in Lebanon. *Journal of Research in Science Teaching, 12*(4), 385-391.

Anila Asghar is assistant professor in the School of Education at Johns Hopkins University. Her work focuses on cognitive and emotional development, science/STEM education reform; science, religion, and culture; and developing scaffolded contexts for faculty development. Correspondence concerning this article can be sent to <anila.asghar@jhu.edu>.

Julie Libarkin is assistant professor of geocognition and tectonics, Department of Geological Sciences, Michigan State University, 206 Natural Science Building, East Lansing, MI 48824. She can be contacted at <libarkin@msu.edu>.

Acknowledgements: We thank all students who graciously participated in this research. Libarkin was partially funded by the National Science Foundation through an Assessment of Student Achievement in Undergraduate Education grant during this project (DUE-0127765). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Aidin Amirshokoohi

Elementary Pre-service Teachers' Environmental Literacy and Views Toward Science, Technology, and Society (STS) Issues

The study explored elementary pre-service teachers' attitudes toward environmental and STS issues, their levels of environmental literacy and knowledge about STS, and their views about teaching environmental and STS issues.

A quick glance at our modern world reveals a deep interrelationship between science, technology, and society. Science and technology are increasingly influencing numerous aspects of contemporary life and are, in turn, affected by societal values and norms. In fact, it is estimated that more than 90% of all current societal issues are grounded in science and technology (Yager, 1987). Hickman, Patrick, and Bybee argue, "The success of individuals and their society is tied to the quality of their choice, which varies with the knowledge and cognitive skills of decision makers." Furthermore, he argues that the success of democracies hinge upon the "ability of citizens to think effectively about developments in science and technology and their effects on the world" (1987, p. 5). It is therefore imperative that citizens understand the interconnections between science, technology, and society and take an active and responsible role in the decision making processes related to the social application of science and technology. Many individuals remain poorly equipped to deal with multifaceted societal issues that are intertwined with science and technology (Cheek, 1992).

Achieving scientific literacy involves educating students about complex social issues and their underlying scientific and technological principles.

Cognizant of the urgent need for scientific literacy in various arenas including the workforce, scientific literacy for all students has become the centerpiece of science education reform movements for the past several decades and has been touted by major reform documents such as the National Science Education Standards (National Research Council [NRC], 1996). Boyer argued that there is a definite place in the core curriculum for the interconnections between science, technology, and society "because these relationships are among the most important ideas, experiences and traditions common to all of us" (1983, p. 302). Achieving scientific literacy involves educating students about complex social issues and their underlying scientific and technological principles. Hence, learning science in its social context is vital to the success of science education reform.

The Science-Technology-Society (STS) movement flourished in the late seventies and early eighties in an effort to tackle the societal concerns of the time which demanded, as Hofstein and Yager argued, "a different kind of science curriculum" (1982, p. 540). Such issues as overpopulation, various types of pollution, dangers of nuclear proliferation, and shortage of water and other natural resources continue to cause concern and stir debates on all sides. The STS movement aimed to address the need to develop a scientific literate society by providing students with real-world connections between

the classroom and society and a richer understanding of societal issues whose root causes or solutions can be found in science and technology. STS curricula have been designed to help students develop skills that will enable them to be responsible citizens who are able to make educated and well-informed decisions.

Not surprisingly, in 1982, the NSTA position statement called for STS as a new emphasis in K-12 science education and recommended dedicating 15-20% of science instruction to STS issues (Yager & Roy, 1993). The STS framework is based on an interdisciplinary constructivist philosophy that promotes the genuine and active engagement of students in the learning process. According to Yager, the process should "give the students practice in identifying potential problems, collecting data with regard to the problem, considering alternative solutions, and considering the consequences based on a particular decision" (1990, pp. 198-200). The STS approach allows the development of particular skills needed to address a wide range of social and technological endeavors (Bybee, 1987) and, consequently, promotes social responsibility and active engagement (Aikenhead, 1984). The STS curriculum focuses on the reciprocal relationships between social, political, and cultural values and scientific research and technological innovations.

The aforementioned complex interactions between science, technology, and society have generated numerous societal issues which are the foci of the STS curriculum framework. Examples of STS issues include pollution, deforestation, global warming, energy depletion, genetic engineering, stem cell research,

The STS framework is based on an interdisciplinary constructivist philosophy that promotes the genuine and active engagement of students in the learning process.

biological and chemical warfare, and nuclear and toxic waste disposal. It is important to note that all environmental issues are STS issues, but not all STS issues are environmental. In fact, the STS curricula have been significantly influenced by the Environmental Education (EE) curricula, which also aim to produce a responsible citizenry. The goal of EE is environmental literacy, which is defined by Roth as "essentially the degree of our capacity to perceive and interpret the relative health of environmental systems and to take appropriate action to maintain, restore, or improve the health of those systems" (1992, p. 14). This echoes an earlier and well documented claim by Stapp (1969) that the goal of environmental education is to create a citizenry that is well-informed about the biophysical environment and its related problems, conscious of ways to help solve those issues, and motivated to work toward the resolution of these issues. The various definitions of the term environmental literacy include several overlapping and related dimensions: environmental sensitivity, knowledge, skills, attitudes and values, personal investment and responsibility, and active invo1vement (Disinger & Roth, 1992). The aim of STS education is to enhance students' scientific and environmental literacy in an effort to bring about changes in personal perception and public policies and, ultimately, to bring about the resolution of STS issues.

Prior studies have suggested a multitude of benefits brought about through STS education, including the development and promotion of scientific "habits of mind" (Hungerford & Volk, 1990; Roth, 1992), positive attitude toward science, increased interest in learning (Yager & Penick, 1991), decision making skills, creativity, and overall science process skills (National Science Teachers Association [NSTA], 1990; Yager, 1989). STS curriculum components encourage students to gain experience in "identifying potential problems, collecting data with regards to the problem, considering alternative solutions, and the consequences based on a particular decision" (Yager, 1990). Similarly, Zoller argues that the STS-oriented approach fosters critical thinking as students become "experts at problem solving, asking questioning, and drawing conclusions based on their interpretation of the societal events" (1992, pp. 289-290). Brunkhorst and Yager (1990) also suggest that STS programs promote higher order thinking skills. STS issues are motivating and thought provoking, and STS education provides students opportunities to 1) interact with their peers, teachers, school, and community, 2) apply their knowledge to real world situations (Yager, Mackinnu, & Blunk, 1992), and 3) experience science outside classroom boundaries.

However, despite the numerous benefits STS education offers, it has, regrettably, not been as widely embraced as originally anticipated. One possible explanation is that teachers are inadequately trained to address science in its social context, and this is due, in part, to the fact

that STS education simply does not "fit with the way education is now structured and presented" (Hausbeck, Milbrath, & Enright, 1992, pp. 32-33). Hence, a logical precursor to the implementation of STS education is to better prepare teachers to adopt this type of science instruction. Teachers are crucial change agents whose classroom practices are immensely influenced by their beliefs (Rubba, 1991). Furthermore, teachers' beliefs have been demonstrated to extensively impact the success of science education reforms in the classroom. Therefore, as Rubba (1991) argues, the development and implementation of an STS curriculum necessitates compatibility between teachers' beliefs and the goals of STS education. Consequently, it is imperative that teacher training involves ample opportunity for teachers to examine their beliefs and confront possible inconsistencies in their beliefs.

Although there have been a number of studies that have explored secondary in-service and pre-service teachers' beliefs about STS education, levels of environmental literacy, attitudes toward teaching STS or environmental issues, and/or the impact of teacher education programs on teacher beliefs and attitudes, studies involving elementary pre-service teachers are extremely scarce. In an effort to begin to bridge this gap in the literature, the current study was initiated to serve as a descriptive exploratory field test for a subsequent study that intends to examine the impact that a STS-oriented science methods course has on pre-service elementary teachers' level of environmental literacy, their views and perceptions toward STS issues, and the instruction of such issues in their classrooms. The aim of the current study was to gain a better understanding of elementary pre-service teachers' attitudes toward environmental and STS issues, their levels of environmental literacy and knowledge about STS, and their views about teaching environmental and STS issues to their prospective students.

Methodology

Sample

This pilot study, which took place at a large Midwestern university in the spring of 2006, is intended to serve as a prelude to a larger study that focuses on the impact of a STS-oriented science methods course on elementary pre-service teachers' environmental literacy and views and perceptions regarding STS issues and instructions. Therefore, it was the intention of the current study to explore the abovementioned factors in this particular population of pre-service teachers to ascertain whether such intervention was necessary. The sample consisted of two sections of the elementary science methods course (n = 41) that were conveniently selected based on the author's and course instructors' schedules. This course was the only science methods course in which elementary education candidates were required to enroll as part of the program. The prerequisite to this course was an introductory science content course especially designed for elementary education majors focusing on scientific inquiry and basic elementary science concepts. Some students had also enrolled in one or both of the science content courses (physics and biology) required for elementary education candidates concurrent with their enrollment in the science methods course. The demographic survey at the beginning of one of the two instruments that were administered revealed several possibly relevant features of this group (summarized in Figures 1-3). The majority of the participants were female (85%), Caucasian students from suburban communities who were completing their last year in the program.

Figure 1: Participants' Residential Communities

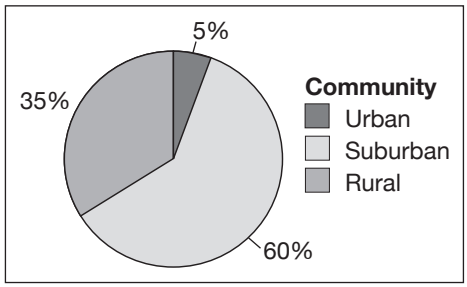


Figure 2: Participants' Racial Backgrounds

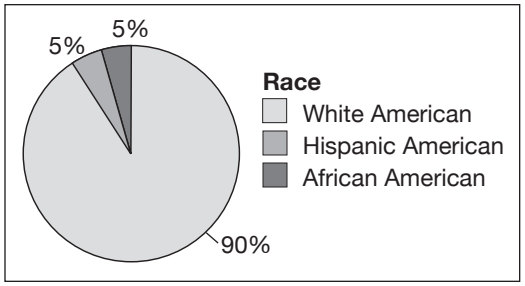
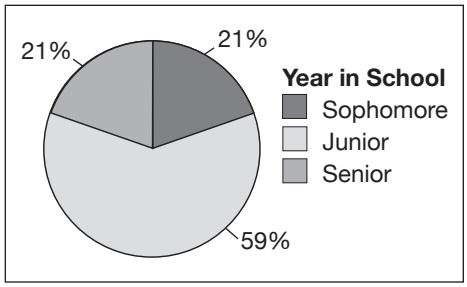


Figure 3: Participants' Year in School



Data Collection

Data collection consisted of the administration of two separate instruments toward the end of the spring semester. The first survey, the Environmental Literacy Instrument

(ELI-7th edition), was developed by Wilke, Hungerford, Volk, and Bluhm (1995) and measures seven subscales: Issue Familiarity, Perceived Knowledge, Perceived Skills, Personal Action History, Issue Identification, Issue Analysis, and Action Plan. Face and content validity for the ELI were established by a national panel of 19 science/environmental education professionals, including university professors, teacher educators, and non-formal professionals (Ngwidibah, 1997). The instruments' reliability measures, which were not reported in the original study, were determined in this current study and will be reported in the results section. The ELI begins with a section that deals with demographics and familial/personal environmental sensitivity indicators.

The main body of the ELI consists of two major tests labeled Test One and Test Two, respectively, plus 12 subsections (see Table 1).

The second instrument, The Perception of STS Issues (PSTSI), was developed by Jamuluddin (unpublished dissertation, 1990) and revised by Ngwidibah (1997). It measures participants' perception of STS issues and the teaching of such issues to elementary students. The Perception of STS Issues (PSTSI) instrument consists of two parts, each containing four questions that utilize a five point Likert Scale (0 = "to no extent"; 4 = "to a great extent"). Part 1, "You and STS Issues", asks participants to answer questions related to: 1) their views regarding the importance of understanding STS issues, 2) their personal interest in understanding STS issues, 3) their perceptions of their own skills to investigate and evaluate STS issues, and 4) their perceptions of their own skills to help resolve STS issues. Questions in part 2, "Teaching STS to Children in Elementary School", focus on: 1) participants' belief about the importance of elementary school students' understanding of STS issues, 2) their willingness to teach elementary school students to understand STS issues, 3) perceptions about their ability to teach elementary school students to investigate and evaluate STS issues, and 4) perceptions about their ability to teach them how to resolve STS issues.

The reliability of the scoring protocol had been established using the inter-rater reliability method of three-way scoring procedure based on random selection of 10 responses from Part I. The original study reported Pearson Correlation Coefficients of .98, .97, and .92 (Ngwidibah, 1997) which correspond to coefficient of determination (r^2) values of .96, .94, and .85 that indicate high level of consistency among the three scorers.

However, despite the numerous benefits STS education offers, it has, regrettably, not been as widely embraced as originally anticipated.

Data Analysis and Results

The different sections of the instruments were scored by the author and a second scorer based on the rubric provided in the instruments (Appendix A). The Cronbach's Alpha for the entire ELI instrument was determined to be 0.75. Because each of the subscales consists of only one item, the reliability for individual subscales cannot be reported. However, since tests 2.1 and 2.2 each consists of three subscales Cronbach's Alpha scores are reported for these two tests as 0.59 and 0.63 respectively. Descriptive statistical analysis, including measures of central tendency, measures of dispersion, and frequency distribution, for the two total scores (ELI & STS totals) were performed and will be discussed in this section.

Table 2 summarizes the results of the ELI. The maximum and minimum scores, the mean, and standard deviation for each of the seven subsections as well as the total score for the entire survey are reported. Due to the lack of similar studies in the literature, the maximum

Table 1: Environmental Literacy Instrument (ELI) sections and scoring rubric

Title	Focus/ Measured Variables	Scale	Maximum Score
Test 1: The Issues with Which I am Familiar [Issue Awareness]	Familiarity with environmental/STS issues	0-4 point each	24
Test 2.1 How I feel About Things and What I Do about Things	Perceived knowledge	0-4	16
	Perceived skill	0-4	16
	Self-reported actions	0-5	140
Test 2.2 [Issue Analysis & Citizenship Action]	Issue Identification	0-6	6
	Issue Analysis	0-16	16
	Citizen Action Skills	0-20	20

Table 2: Results of the ELI Instrument

	Maximum possible score	Minimum score	Maximum score	Mean	Standard deviation
Issue Familiarity	24	0	16	3.27	3.362
Perceived Knowledge	15	0	10	5.22	2.253
Perceived Skills	15	0	11	4.78	2.715
Personal Action	140	5	59	25.53	12.743
History Issue	15	0	5	4.29	2.251
Identification Issue	15	0	15	8.24	7.003
Analysis Action Plan	20	0	20	11.75	5.350
Grand Total for ELI	238	16	100	63.68	19.150

possible scores for each section and the entire instrument served as a point of reference in the interpretation of the data. The data indicate low scores on all subsections and the overall instrument. The mean scores were incredibly low for the first five sections: Issue Familiarity (M=3.27, SD=3.362), Perceived Knowledge (M=5.22, SD=2.25), Perceived Skills (M=4.78, SD= 2.715), Personal Action History (M=25.53, SD=12.74), and Issue Identification (M=4.29, SD=2.25) and failed to reach even one half of the total possible score for each respective section. Participants' mean scores on the last two sections, Issue Analysis (M= 8.24, SD=7.00), and Action Plan (M=11.75, SD=5.35) fared better than the aforementioned ones and slightly exceeded 50% of the total possible score for these sections. These sections differed from the first five in that they either provided participants choices to select from or background stories that they could read and extract necessary information from. These differences might serve as a possible explanation for the higher scores on these sections than the other sections. The mean grand total score for the ELI instrument (M=63.68, SD=19.15) failed to reach even one third of the total possible score. The maximum score of 100, obtained by only one participant, was still considerably lower than the total possible score of 230.

Figure 4 provides a visual of the frequency distribution of the Grand Total ELI scores. Most participants scored between 60 and 90 on this instrument. These scores indicate that the results are skewed to the left.

Figure 4: Frequency Distribution of the Grand Total ELI Scores

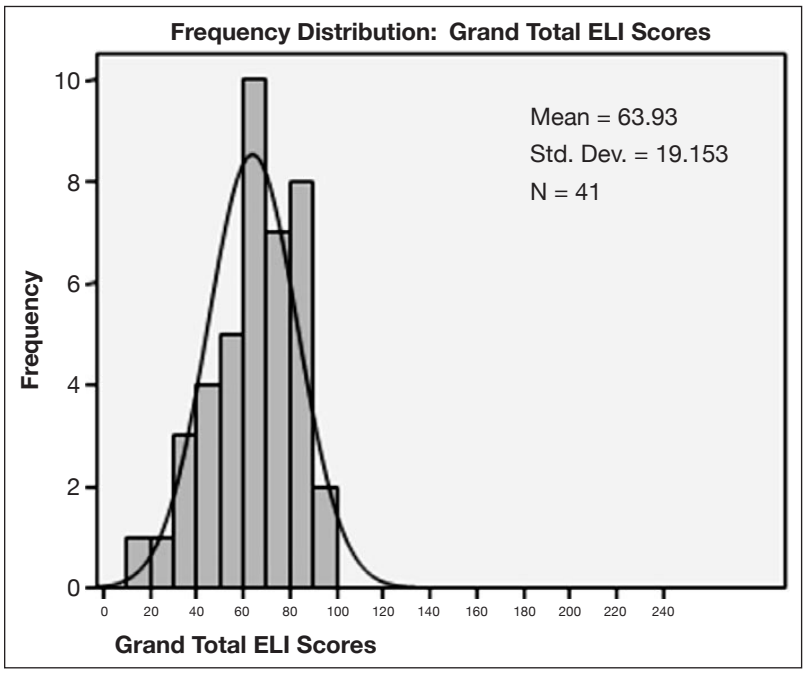


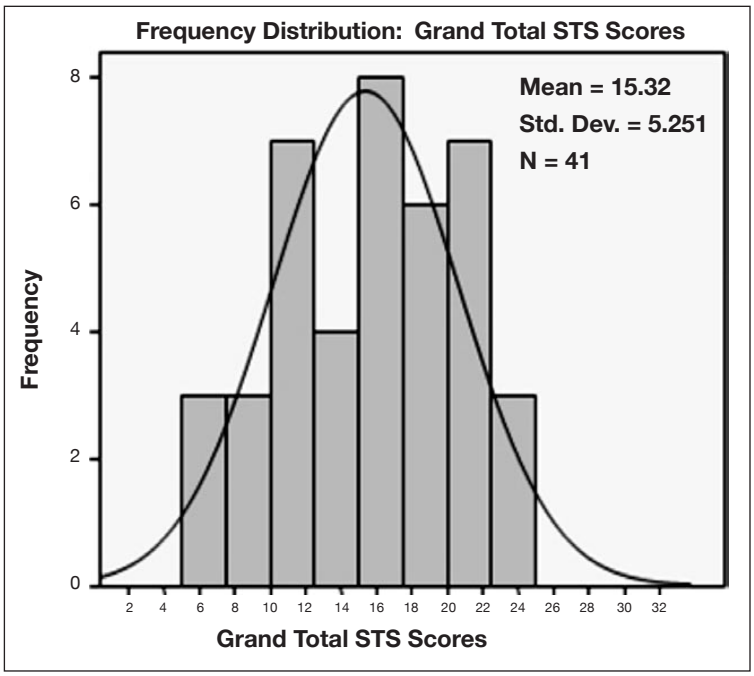
Table 3 indicates the results of the PSTS survey. Similar to the ELI instrument, the data for this instrument were also interpreted based on the maximum possible scores. However, a quick glance at the data from the two instruments reveals that the participant scores on this instrument were comparatively better than the ELI scores. For example, the grand total mean score (M=15.32, SD=5.25) for this instrument was found to approximate 50% of the total possible score of 32. Figure 5 shows the frequency distribution of the Grand Total scores for the PSTS instrument. The frequency distribution of participants' scores on this instrument is not as skewed as the other instrument. The majority of the scores fell between 10 and 22. Their overall score for the perceptions of teaching STS in the classroom section [M=8.85 (55%)] was higher than the overall score for the section on personal views of STS [M=6.46 (40%)], which indicates that these teacher candidates had a more positive attitude toward *teaching* STS issues than toward personal awareness of these issues. The mean scores for the sections dealing with their views about the importance of understanding STS issues (M=2.27) and teaching students about

Table 3: Perceptions of STS & STS Teaching

	Maximum possible score	Minimum score	Maximum score	Mean	Std. Deviation
STS Q1: Importance of Understanding STS issues	4	0	4	2.27	.923
STS Q2: Personal interest in understanding STS issues	4	0	3	1.56	.867
STS Q3: Skills to investigate & evaluate STS issues	4	0	3	1.39	.919
STS Q4: Skills to resolve STS issues	4	0	3	1.24	.734
STS Total	16	1	10	6.46	2.570
STS teaching Q1: Importance of knowledge of STS for elementary students	4	0	4	2.39	.972
STS teaching Q2: Willingness to teach STS to elementary students	4	0	4	2.44	.976
STS teaching Q3: Ability to teach elementary students to evaluate STS issues	4	0	4	2.10	.944
STS teaching Q4: Ability to teach elementary students to resolve STS issues	4	0	4	1.93	.959
STS teaching Total	16	0	16	8.85	3.461
Grand Total for STS Instrument	32	5	24	15.32	5.251

STS (M=2.39) as well as their willingness to teach STS (M=2.44) were the highest. Their personal interest in STS (M=1.56), perceptions of their skills to investigate STS issues (M=1.39), and perceptions of their skills to resolve STS issues (M=1.24) were the lowest scores.

Figure 5: Frequency Distribution of Grand Total PSTS Scores



Discussion and Implication

The results of this study indicated low levels of environmental literacy among this sample of elementary pre-service teachers. They were also indicative of low levels of personal interest in STS issues. The slightly better scores on the section of the PSTS instrument dealing with participants' views toward teaching STS in the classroom bode well and provide hope that if equipped with sufficient understanding of STS

and environmental issues and the STS instructional framework, these participants are willing and consider it important to teach their prospective students about these issues. Their low scores on the environmental literacy instrument and the STS section of the PSTS indicate that further training must be provided to augment elementary pre-service teachers' understanding of environmental and STS issues and enhance their own views about such issues. It is evident that their prior beliefs and understanding of such issues are not consistent with STS education reform. Teacher training programs must, therefore, allow prospective elementary teacher candidates the opportunity to critically reflect on their beliefs and knowledge regarding STS/EE issues and STS/EE education. Only when their beliefs are aligned with the STS-oriented framework of teaching science and their levels of environmental and STS literacy are enhanced can we expect prospective elementary education candidates to be willing and able to implement such instruction in their future classrooms.

There remain countless gaps in the literature on pre-service teachers' environmental literacy and views toward STS/EE. This is especially true for elementary pre-service teachers. Further research is necessary in several areas. First, there is a need for replication studies to further explore elementary pre-service teachers' environmental literacy and views toward STS issues and instruction. Second, possible factors affecting pre-service teachers' environmental literacy and beliefs about STS prior to entering teacher education programs must be explored. Third, factors within the teacher training programs that might influence prospective teachers' knowledge and beliefs in these areas must also be explored. The possible impact of STS-oriented science methods courses on teacher candidates' levels of environmental literacy and perceptions of STS demand significant attention. Furthermore, within such methods courses, factors that may lead to possible changes in participant beliefs and understanding should be explored to allow for replication of these types of courses in other programs. The aforementioned questions will be examined in the subsequent larger study following the pilot study. Finally, studies such as the current pilot study should be replicated in other geographical regions of the country with various groups of pre-service teachers to explore whether the same trends occur with other populations. Additionally, replication might enable science teacher educators to identify possible factors that influence teacher candidates' prior understanding and beliefs.

References

Aikenhead, G. S. (1984). Teacher decision making: The case of Prairie High. *Journal of Research in Science Teaching*, 21 (2), 167-186.

Boyer, E.L. (1983). High School: A Report on Secondary Education in America. New York: Harper & Row. 302-303.

Brunkhorst, H.K. and Yager, R.E. (1990). Beneficiaries or victims. *School Science and Mathematics*, 90 (1), 61-69

Bybee, R. W. (1987). Teaching about science-technology-society (STS): Views of science educators in the United States. *School Science and Mathematics*, 87 (4), 274-285.

Cheek, Dennis W. (1992). *Thinking constructively about science, technology and society education*. Albany: SUNY Press.

Disinger, J. F. and Roth, C. E. (1992). Environmental Literacy. Columbus, OH: Education Resources Information Center/Center for Science, *Mathematics and Environmental Education*.

Hausbeck KW, Milbrath LW, Enright SM. 1992. Environmental knowledge, awareness and concern among 11th grade students: New York State. *Journal of Environmental Education* 24(1):27-34.

Hickman, F., Patrick, J., & Bybee, R. (1987). Science/technology/society: A framework for curriculum reform in secondary school science and social studies. Boulder, CO: Social Science Education Consortium, Inc.

Hofstein, A and Yager, R. (1982). Societal Issues as Organizers for Science Education in the 80s, *School Science and Mathematics*, Volume 82, pp.539-547.

Hungerford, H. & Volk, T. (1990). Changing learner behavior through environmental education. *Journal of Environmental Education 21* (3): 8-21.

Jamuluddin, S. (1990). The effects of an STS course on selected variables associated with STS issue instruction and citizenship behavior in preservice elementary teachers. Unpublished doctoral dissertation, Southern Illinois University, Carbondale, IL.

National Research Council. (1996). *National science education standards*. Washington, D.C.: National Academy Press.

National Science Teachers Association. (1990). Science teachers speak out: The NSTA lead paper on science and technology education for the 21st century. Washington, D.C.: National Science Teachers Association.

Ngwidibah, L. N. (1997). The effects of an Environmental/Science-Technology-Societal issues curriculum on preservice teachers. Unpublished doctoral dissertation, Southern Illinois University, Carbondale, IL.

Roth, Charles E. (1992). Environmental literacy: Its roots, evolution, and directions in the 1990s. Columbus, OH: ERIC/CSMEE. Roth, Charles E. (1968). On the road to conservation. Massachusetts Audubon, June 1968, pp. 38-41.

Rubba, P.A. (1991). Integrating STS into school science and teacher education: Beyond awareness. *Theory into Practice*, 30(4), 303-308.

Stapp, W.B. (1969). The concept of environmental education. *Journal of Environmental Education* 1(3): 31-36.

Wilke, R., Hungerford, H., Volk, T., & Bluhm, W. (1995). *Middle School Environmental Literacy Instrument*. Unpublished Instrument

Yager, R. E. (1987). STS science teaching emphasizes problem solving. *Curriculum Review*, 38-41.

Yager, R. (1989). What science should contribute to cultural literacy. *The Clearing House*, 62, 297-302.

Yager, R. E. (1990). The science/technology/society movement in the United states: Its origin, evolution, and rationale. *Social education, 54* (4), 198-200.

Yager, R.E., Mackinnu, Blunck, S.M. (1992). Science/technology/society as a reform of science in the elementary school. *Journal of Elementary Science education*, 4(1), 1-13

Yager, R.E. & Penick, J.E. (1991). Using science as a connector for the reformed school curriculum. NAAP Bulletin, 76 (547), 56-68

Yager, R. E., & Roy, R. (1993). STS: Most pervasive and most radical of reform approaches to "science" Education. In R.E. Yager (Ed.) The science, technology, society movement (pp. 7-13). National Science Teachers Association.

Zoller, U. (1992). The Technology/Education Interface: STES Education for All. *Canadian Journal of Education*, 17(1), 86-91

Aidin Amirshokoohi is an Assistant Professor of Science Education at Fairfield University in Fairfield, CT where he also serves as the science education program coordinator. He has been involved in leading several grant-sponsored professional development workshops for in-service science teachers. During the past few years, he has co-authored several papers and presented at NARST, ASTE, and SSMA national conferences. Correspondence concerning this article can be sent to <aamirshokoohi@fairfield.edu>.

Ronald K. Atwood, John E. Christopher Rebecca K. Combs, Elizabeth E. Roland

In-Service Elementary Teachers' Understanding of Magnetism Concepts Before and After Non-Traditional Instruction

The authors provide a descriptive study of in-service elementary teachers' understanding of magnetism concepts and confidence in their understanding of those concepts before and after non-traditional instruction that utilizes instructional activities from Physics by Inquiry.

Introduction

Magnetism is a topic frequently studied in elementary schools (Toleman, 1998). Since magnetism is a popular topic and is included in national science education standards (American Association for the Advancement of Science, 2003; National Research Council, 1996), it might be assumed that elementary teachers have a good understanding of this topic and that elementary students develop a good understanding of fundamental magnetism concepts. Unfortunately, evidence suggests that magnetism concepts are poorly understood across a broad range of potential learners (Atwood, Christopher & McNall, 2007; Constantinou, Raftopoulos, & Spanoudis, 2001; Finley, 1986; Hickey & Schibeci, 1999). The lack of successful teaching and learning of magnetism concepts that occurs at the elementary level may be partly due to deficiencies in elementary science textbooks (Barrow, 1990) for elementary students and elementary science methods and materials textbooks (Barrow, 2000) for teachers. However, ineffective science content courses in teacher preparation are likely to be a much larger problem (McDermott, 1991; McDermott, Heron, Shaffer, & Stetzer, 2006). There is a clearly identified need to improve instruction on magnetism, and elementary science teacher education is a logical place to focus. A study of pre-service (Atwood & Christopher, 2007) teachers has revealed a poor understanding of basic magnetism concepts, and traditional survey science courses may be doing little to improve that situation. The documentation of inadequate understanding of standards-based magnetism concepts by elementary students and teachers is an important start to understanding the nature and magnitude of this problem, but it is also necessary to address the lack of conceptual understanding of elementary teachers.

The idea that understanding is domain-specific is central to thinking about conceptual understanding and change.

Theoretical Framework

The following considerations were used to identify characteristics of instruction likely to be associated with the desired impact:

1. The goal of the instruction is to facilitate teachers' construction of conceptual understanding of basic magnetism concepts.
2. Traditional instruction has failed to result in the desired understanding, so it is unsuitable for the study (McDermott, 1991; McDermott, Heron, Shaffer, & Stetzer, 2006).
3. Minimally guided, non-traditional instruction has been strongly criticized recently and is unlikely to result in the desired understanding (Kirschner, Sweller, & Clark, 2006; Mayer, 2004).

4. Non-traditional investigative instruction that is judiciously structured and scaffolded and consistent with the intentional conceptual change literature has shown great promise and should be utilized in the study (Beeth, 1998; Niaz, 1995; Nussbaum & Novick, 1982; Vosniadou, 2003, 2007).

For some time, the science education community has shown considerable support for teaching for understanding (American Association for the Advancement of Science, 2003; Gallagher, 2000; Gardner & Boix-Mansilla, 1994; National Research Council, 1996; Prawat, 1989; Wildey & Wallace, 1995). During roughly the same period, it has been well documented that diverse populations of children and adults lack a scientific understanding of many fundamental science concepts across the biological, earth, and physical sciences (Atwood & Christopher, 2007; Bar, 1989; Barman & Griffiths, 1995; Baxter, 1989; Brody & Koch, 1990; Driver, Guesne, & Tiberghien, 1985; Duit, 1984, 2004; Duit & Treagust, 1995; Krall, Christopher, & Atwood, 2009; Osborn & Cosgrove, 1983; Schoon, 1992; Trundle, Atwood & Christopher, 2002).

The pervasive lack of conceptual understanding has been partially attributed to the failure of traditional instruction, a term that seems to be a broad umbrella for a variety of presentation modes. Textbooks and lectures have historically been the most popular modes for presenting information, but these methods typically do not use the collection and analysis of data as a basis for generating explanations. Although computers and other technology are increasingly used as presentation modes, as well as for more creative purposes, the basic approaches behind the instructional methods often remain largely unaltered.

Regardless of when they are formed, non-scientific conceptions can be organized into a durable, theory-like framework that has explanatory capacity and is resistant to change.

Minimally guided, non-traditional instructional approaches have been utilized as alternatives to traditional presentation modes, but their effectiveness has recently been strongly criticized (Kirschner, Sweller, & Clark, 2006; Mayer, 2004). Critics have tended to paint all efforts to utilize discovery, constructivist, inquiry-based, and problem-solving approaches with the same broad brush. They have argued that minimally guided instructional approaches fail, because they place a heavy burden on students' cognitive processing that prevents students from processing novel information.

However, non-traditional instruction consistent with conceptual change theory is judiciously structured and scaffolded yet still investigative in nature (Beeth, 1998; Champagne, Gunstone, & Klopfer, 1985; Cosgrove & Osborn, 1985; Hewson & Hewson, 1983, 1984; Jackson, Dukerich, & Hestenes, 2008; Niaz, 1995; Nussbaum & Novick, 1982; Osborn & Wittrock, 1983; Smith, Blakessie, & Anderson, 1993; Vosniadou, 2003; Weaver, 1998). The work of Vosniadou (1991, 2002, 2003, 2007) has been particularly useful in exploring the ways in which conceptual change theory should inform instruction, perhaps because her views have been influenced by personal research on non-scientific conceptions (Vosniadou, 2003; Vosniadou & Brewer, 1994).

The idea that understanding is domain-specific is central to thinking about conceptual understanding and change (Carey, 1985). Further, although understanding is constructed internally by an individual, the process may be influenced externally by a number of factors. Socio-cultural influences, such as formal schooling, can be among the most important external factors (Hatano & Inagaki, 2003). Conceptions formed prior to formal schooling tend to be naïve and non-scientific. Regardless of when they are formed, non-scientific conceptions can be organized into a durable, theory-like framework that has explanatory capacity and is resistant to change. Becoming metacognitively aware of how one's understanding compares with the accepted scientific understanding and being motivated to adopt the scientific understanding, seem crucial for the radical restructuring that is sometimes required (Vosniadou, 2003, 2007). To achieve intentional conceptual change, it is necessary to go beyond active learning. Sinatra and Pintrich described this process as "the goal-directed and conscious initiation and regulation of cognitive, metacognitive and motivational processes to bring about a change in knowledge" (2003, p. 6). Although instruction has a critical role to play in helping students to achieve the conceptual change that must take place in order to gain understanding of fundamental science concepts, it seems highly unlikely that either a traditional presentation mode of instruction or minimally guided

non-traditional modes of instruction will create the conditions needed for intentional conceptual change to occur.

The Instruction

McDermott (1991) and others have described the inadequacies of traditional presentation-mode science survey instruction for teachers (McDermott et al., 2006). McDermott has led a group in the development of instructional materials, called *Physics by Inquiry* (McDermott, 1996), that appear to be strongly aligned with the instructional characteristics needed to promote intentional conceptual change. Further, use of *Physics by Inquiry* instructional materials has been associated with sharply increased performance among pre-service elementary teachers studying several other science topics, such as moon phases (Trundle, Atwood & Christopher, 2002) and force and motion concepts (Arts, 2006). Comparable positive results have been found for in-service middle school teachers' study of light phenomena (Atwood, Christopher & McNall, 2006). The *Physics by Inquiry* materials are structured to encourage students to take responsibility for their own active learning. This promotes metacognitive processing, because students must compare the explanations they have previously held in their conceptual frameworks with explanations that explain the data they generate through investigations. In this way, *Physics by Inquiry* instructional materials about magnetism support both empirical data and intentional conceptual change theory.

In the present study, instruction on magnetism was provided as part of a one-week physical science institute for in-service elementary teachers.

Although instruction has a critical role to play in helping students to achieve the conceptual change that must take place in order to gain understanding of fundamental science concepts, it seems highly unlikely that either a traditional presentation mode of instruction or minimally guided non-traditional modes of instruction will create the conditions needed for intentional conceptual change to occur.

Approximately five hours of the 30 available for instruction during the week were devoted to magnetism, and the remainder was used to address other physical science topics. *Physics by Inquiry* was the source of activity ideas and instructional strategies. It should be noted that the investigators selected only a fraction of the activities and materials available in the magnetism section of *Physics by Inquiry*. Due to the time limitation, activities and materials judged to be most fundamental to the schools' K-4 content standards were selected. While participating in these activities, teachers work in small groups to complete investigations, make and discuss observations, and arrive at conclusions.

In one activity, participants were given two bar magnets (identified as such) and a tray of objects made from a variety of materials. They were asked to explore the character of any interaction between the magnets and between the magnets and other objects. Then, they were led to classify the objects into three categories on the basis of the observed magnetic interactions. The three categories were later identified as magnets, ferromagnetic materials, and non-magnetic materials. Participants prepared an evidence-based procedure for confidently determining whether a magnet was included among a group of objects. Next, they studied in more detail the interactions between the parts (ends and middle) of two bar magnets with each other and also with similarly shaped ferromagnetic materials. This led to the introduction of the idea of magnetic poles, and participants then discussed methods of finding the poles and identifying them as north-seeking or south-seeking. The teachers then participated in activities that allowed them to locate the positions of the two poles on a variety of familiar magnets. Next, they observed and described the behavior of magnetic compasses. Participants were also given a magnetic model of the earth, and they discussed how magnetic compass needles behave in relation to the model. At this point, each group was issued a paper clip and a ruler and instructed to use those materials to develop a procedure for comparing the strength of two magnets. The groups then used their procedure to order the strength of the poles of a number of magnets that had different sizes and shapes. Finally, the small groups explored the analogy between a bar magnet and a stack of flat, rectangular "refrigerator" magnets (with the poles on the faces). They identified the pole locations and types of poles for a stack, and they compared the behavior of stacks of varying sizes.

In summary, this non-traditional, guided inquiry instruction frequently engaged the in-service teachers in

making systematic observations and engaging in interpretive discussions of their own observations. They also prepared responses to three *checks,* which were written conclusions based on previous observations and responses to challenging application questions. Each check was completed in written form by each individual, discussed in a small group, and then defended during a discussion with an instructor. This constructivist design encouraged participants to maintain a high degree of awareness of their own thinking and understanding as they mentally processed a steady inflow of observations and made conjectures. As suggested previously, instruction with these characteristics has a high potential for facilitating intentional conceptual change (Bereiter & Scardamalia, 1989; Hennessy, 2003).

Research Questions

The central research question for this descriptive study is: How does the conceptual understanding of selected magnetism concepts compare before and after in-service elementary teachers complete non-traditional instruction from *Physics by Inquiry*? A secondary research question is: How does the confidence that in-service elementary teachers have in responding to assessment tasks on magnetism compare before and after completing non-traditional instruction from *Physics by Inquiry*?

Procedures and methodology

The 18 elementary teachers in the non-random sample self-selected into a one-week physical science institute. The teachers were from four rural school districts in central Appalachia. Oliver (2007) has described the difficulty of defining rural in an era of greatly reduced isolation due to the internet and interstate highways. We use the rural school district description here for communities that are heavily dependent on agriculture in their economies and lack a town of more than 5,000. In addition to the approximately five hours of instruction provided on magnetism concepts, during the remaining 25 hours of instructional time, physical properties, light phenomena and force and motion were also addressed. Considerably more instructional time was devoted to light phenomena and force and motion than magnetism, because it was assumed that magnetism was more likely to have been regularly taught by elementary teachers. Given that, it was reasoned teachers were more likely to have learned what they needed about magnetism than light phenomena and force and motion. Additionally, prior to determining the institute topics, instructional supervisors in the region expressed the view that local teachers were better prepared to teach magnetism than light phenomena and force and motion.

Five multiple-choice tasks with popular non-scientific conceptions embedded in the distracter options (Hestenes, Wells, & Swackhamer, 1992) were a major source of data in addressing the central research question. In addition, teachers were asked to provide a brief, written justification or explanation for each multiple-choice selection. These supporting statements were expected to provide insight into the reasoning used to make the multiple-choice selections. Finally, the teachers rated their confidence in the correctness of each answer using a five-point scale. The confidence level descriptions and corresponding numerical ratings were as follows:

1. Highly confident
2. Somewhat confident
3. Neutral in confidence
4. Somewhat lacking confidence
5. No confidence

The confidence data were used to address the secondary research question. Since magnetism is a popular topic in elementary schools, it was thought that the teachers might be confident in responding to magnetism questions. The teachers were given instructions for generating and using a code on all assessment forms to insure anonymity and minimize anxiety while allowing pre-test and post-test matching of individuals' scores. The assessment tasks were administered, along with tasks addressing the other physical science topics, during the beginning and closing hour of the institute.

If teachers have not been adequately prepared to teach fundamental concepts about magnets and the behavior of magnets, there are important implications for both pre-service and in-service teacher education.

Results and Discussion

The results are presented and discussed by multiple-choice task. A representative sample of participants' explanations of multiple-choice option selections is included in the discussion. The self-reported confidence rating is located in parentheses immediately after each explanation. For each task, the multiple-choice pre-test data that showed the frequency with which each

option was selected were cross-classified by the post-test data. This arrangement facilitates the analysis of changes in multiple-choice responses from pre-test to post-test. In order to give the reader an understanding of the advantages of this arrangement, Table 1 shows the data arranged in this way for Task 1. In the interest of conserving space, similar tables that were used in the analysis of Tasks 2-5 are not included here, but interesting changes in responses from pre-test to post-test are described in the text. Table 2 provides a summary of the data corresponding to the multiple-choice and supporting explanation or justification for each of the five tasks. The confidence data for all five tasks have been placed in Table 3 to help communicate the shifts in confidence that occurred.

Table 1: Task 1 Comparison of pre-/post-test multiple-choice selections by in-service teachers.

Results reported as frequencies, except as noted

		Post						Pre-Totals	Pre-Totals
		A*	**B**	**C**	**D**	**E**	**Omit**	**as *f***	**as %**
	A*	7	0	0	0	0	0	7	38.9
	B	0	0	0	0	0	0	0	0.0
	C	4	0	2	0	0	0	6	33.3
	D	3	0	0	0	0	0	3	16.7
Pre-	**E**	2	0	0	0	0	0	2	11.1
	Omit	0	0	0	0	0	0	0	0.0
Post- Totals as *f*		16	0	2	0	0	0	18	
Post- Totals as %		88.9	0.0	11.1	0.0	0.0	0.0		100

Note. The '*' marks the correct letter choice for the task.

Using magnets to attach a variety of objects to refrigerator doors is a common practice in homes. Task 1 provided an opportunity to show an understanding of the science involved. Table 1 and Table 2 provide a summary of multiple-choice results for Task 1, and Table 3 includes the confidence levels associated with each response.

Teachers who understand that very few materials, including iron, are ferromagnetic and will interact with a magnet can demonstrate that understanding by choosing option A. Table 1 reveals that only 7 of 18 (39%) participants selected A on the pre-test, but 16 of 18 (89%) did so on the post-test. Option C was the most popular distracter option on the pre-test, as it was selected by six (33%) teachers. Only two teachers selected this option on the post-test, and these were the only persons who failed to select the correct response. Refrigerator doors require little effort to open and close, and this may have led to the mistaken conclusion that a lightweight metal, such as aluminum, was used in the doors. However, if the teachers had ever had the experience of testing several different known metals, including iron and aluminum, it seems likely that option A would have been the preferred choice on the pre-test. Because it is known that a popular and strongly held non-scientific conception is that magnets attract most metals (Hickey & Schibeci, 1999), the opportunity to test several labeled metals, including aluminum, and discuss the results was provided during the instruction. It appears that experience was sufficient for four of the six participants who selected C on the pre-test.

Figure 1: Task 1 assesses interactions between magnets and ferromagnetic materials.

1. The most likely reason magnets stick to refrigerator doors is because they are interacting with
   A. iron in the doors.
   B. the plastic or ceramic coating on the doors.
   C. a lightweight metal, such as aluminum, in the doors.
   D. a heavy metal, such as lead, in the doors.
   E. electric charge on the refrigerator doors.

Examination of the pre-test explanations or justifications showed that two of the seven teachers who selected the correct answer simply stated that "magnets stick to iron" (2, 4), two wrote "magnets are attracted to iron" (2, 4), and a fifth explained why the other four options were not correct. Another correct pre-test responder provided no written explanation (3), and still another explained, "magnets have to attract to metal or iron" (3). The latter response appears to represent a false positive, a correct response based on a non-scientific reason. False positives represent an important limitation of forced-choice

Table 2: Pre-test and post-test data showing frequencies and percents of correct multiple-choice (MC) responses and correct multiple-choice responses adequately supported by explanation or justification

	Correct Pre-				Correct Post-			
	MC Only		MC and Support		MC Only		MC and Support	
Task	f	%	f	%	f	%	f	%
1	7	38.9	5	27.8	16	88.9	13	72.2
2	14	77.8	9	50.0	17	94.4	13	72.2
3	5	27.8	1	5.6	15	83.3	10	55.6
4	12	66.7	7	38.9	18	100.0	14	77.8
5	11	61.7	2	11.1	9	50.0	6	33.3
Totals	49	54.4	24	27.8	75	83.3	56	62.2

Table 3: Tasks 1-5 summary of magnetism pre- and post-test confidence for in-service elementary teachers

Response Frequencies from 1, Highly Confident to 5, No Confidence												
	Pre-						Post-					
Task	1	2	3	4	5	Total	1	2	3	4	5	Total
1	0	6	5	6	1	18	13	5	0	0	0	18
2	2	8	4	3	1	18	10	7	1	0	0	18
3	0	3	4	7	4	18	9	4	2	1	2	18
4	1	5	5	2	5	18	14	4	0	0	0	18
5	3	3	6	2	4	18	13	5	0	0	0	18

tasks (Trundle, Atwood & Christopher, 2002). Thus, as shown in Table 2, on the pre-test only five of 18 teachers both selected the correct response and provided a satisfactory statement in support of their selection. The average confidence level for the five participants who selected the correct answer to Task 1 on the pre-test and provided a scientifically correct justification was 2.8, a neutral level.

On the post-test, 16 of 18 teachers (88.9%) selected the correct response. However, only 13 of the 16 who selected the correct response also provided a scientific explanation (see Table 2). These 13 teachers had an average confidence level of 1.1. Of the remaining three correct multiple-choice responders, one stated that "magnets are attracted to metals containing lead" (1), one provided no explanation and selected a confidence level of 1, and one admitted to making a "guess" (1). The first and last of these three participants are assumed to represent false positives, and it is possible that the response with no justifying explanation was also a false positive. The fact that the seven teachers who selected the correct multiple-choice response on the pre-test also did so on the post-test is viewed as a positive result. In fact, Table 1 reveals that all movement from one response option to another was positive, from a non-scientific option to the scientific option.

One of the two teachers who selected option C on the post-test explained, "the door must have some type of metal to attract the magnet" (2). This explanation is essentially unchanged from the pre-test and is interpreted as confirmation of a firmly held non-scientific conception. The second teacher explained, "some metals do not attract magnets" (2) but apparently thought that aluminum does.

In Table 3, note that no teacher selected a confidence level of one on the pre-test and only six (33%) selected a two. In comparison, on the post-test 13 persons selected a confidence level of one and the other five selected a two. Overall performance on Task 1 is much improved on the post-test. This is true for selecting the correct multiple-choice response, providing scientific explanations in support of correct responses, and expressing greater confidence in correct selections and explanations. It is interesting to note that even those who failed to support a correct multiple-choice response with a satisfactory explanation were very confident on the post-test.

Task 2 was used to probe participants' understanding that the needle of a magnetic compass aligns to point approximately geographic north and south.

Figure 2: Task 2 assesses an understanding of the interaction of the Earth's magnetic field with a compass.

2. You may use a magnetic compass to find your way,
   A. since the compass needle will always point in the direction you are facing.
   B. during the day but not during the night.
   C. since the compass needle aligns in a north/south direction.
   D. if there aren't too many trees or mountains nearby.
   E. because compass needles don't move.

As shown in Table 2, 14 of the 18 participants (78%) selected the correct answer, C, on the pre-test and 17 participants (94%) did so on the post-test. Analysis of the pre-test results shows that nine of the 14 who made the correct multiple-choice selection gave an explanation that the compass needle always points north, and four of those nine added correct information about magnetic poles in their explanations. However, two of the remaining five gave no explanation (with confidence levels of 3 and 5), and responses by the other three consisted of "not sure" (4), "I think" (2), and "guess" (4). Eight of the nine teachers who gave a satisfactory explanation in support of a correct multiple-choice selection on the pre-test selected one of the top two confidence levels and the other chose the neutral level (3). A low confidence level paired with no substantive explanation supports suspicion, but does not confirm, that the correct multiple-choice selection resulted from an understanding that is not strongly held, or from guessing. Thirteen of the 17 who selected the correct multiple-choice response on the post-test also provided adequate scientific explanations. Representative explanations include "the needle of the compass is magnetic and the north end of the needle will point to the north pole that has a south magnetic pull" (2) and "the N needle always points to the N geographical pole (S magnetic pole)" (1).

On the pre-test, three persons selected option A for Task 2. Although this may be viewed by persons working in the sciences as a curious choice, it is a perspective not uncommonly encountered by the authors in working with both pre-service and in-service elementary teachers. In the present study, the pre-test explanations provided in support of this choice were: "compasses show you what direction you are traveling in" (4), "based on the true north/south the needle will point the way you are facing" (3), and "a compass tells you the direction you are facing" (2). On the post-test, the three persons who had selected option A on the pre-test joined the 14 persons who selected the correct response on both the pre-test and post-test. All three gave satisfactory supporting explanations on the post-test, including "the needle of a compass is magnetic and the north end of the needle will point to the north pole that has a south magnetic pull" (2). Again, all of the movement from one multiple-choice option to another on Task 2 was positive.

In Table 3 note that participating teachers showed more confidence in their pre-test responses to Task 2 than for any other task, as ten teachers chose levels one or two. Confidence increased in post-test responses with 17 of 18 participants selecting confidence levels one or two. The only person who selected an incorrect response on the post-test provided no explanation for it and chose the neutral confidence level.

Bar magnets are frequently used in elementary classrooms. Task 3 (Figure 3) provided an opportunity for the teachers to consider the validity of five statements about bar magnets and the behavior of bar magnets.

A review of the data in Table 2 shows that pre-test performance on Task 3 was the weakest of the set. Only five participants (25%) selected the correct answer, D. The confidence data in Table 3 for Task 3, consistent with the multiple-choice results, also are the lowest of the set. For the five teachers who selected the correct multiple-choice option on the pre-test, the explanations provided were: "process of elimination" (4), "guess" (3), "guess" (4), "has plus and minus charged ends" (4), and "they will repel if turned correctly" (4). Thus, four of the five statements suggest the correct responses on the multiple-choice task may have been false positives, because only one teacher both selected the correct response and provided a satisfactory supporting explanation. Note also that one confidence level was neutral and the other four were below the neutral level, including for the one teacher whose explanation is consistent with a scientific understanding.

Figure 3: Task 3 assesses the expected understanding of the properties and behavior of bar magnets.

3. A bar magnet
   A. has the strongest magnetic effect in the middle of the bar.
   B. interacts with all metallic objects.
   C. will not influence a magnetic compass.
   D. can repel any other magnet.
   E. interacts with heavy metals like lead, brass, and gold.

Each distracter option was attractive to between 11 and 28% of the sample. Four participants selected option A, which was the idea that a bar magnet has the strongest magnetic effect in the middle of the bar. The explanations included an idea the investigators had

not encountered previously, which was that magnetic fields are "strongest where the N/S come together" (3), as well as several previously encountered explanations such as "a process of elimination, I think" (3), "guess" (5), and "I am not sure; I used an educated guess" (5). Note the neutral to low confidence levels associated with these statements.

On the pre-test results for other Task 3 distracters, four teachers selected option B, which indicates a bar magnet interacts with all metallic objects. The varied supporting statements provided were "B seemed like the only one that could be true" (4), "a magnet interacts with metal no matter what shape" (2), "a bar magnet has two charges, plus and minus" (2), and "?" (4). Two of the three persons who selected option C also apparently had little understanding of the properties of bar magnets. One of the two did not provide an explanation (confidence level 4), and the other wrote, "I guessed" (5). The third person who selected C was somewhat confident and explained that "the opposite poles will react and similar poles will not" (2). The two persons who selected option E wrote "attracts to all metals" (3) and "guess" (5). Note that the former respondent expressed neutral confidence while the latter expressed no confidence. Both Tasks 1 and 3 reveal the attractiveness of the non-scientific conception that magnets attract many different metals.

The post-test selection of the correct response on Task 3 by 15 of 18 (83%) participants is viewed as a very favorable result when compared with the pre-test data. Further, 10 of the 15 also provided a satisfactory supporting statement. The five persons who selected the correct response on the pre-test also did so on the post-test. In addition, three of four teachers who had selected option A, all of the persons who had selected options B or C, and one of the two teachers who had selected option E moved to the correct response on the post-test. Two of the three persons who selected an incorrect post-test response had made a different incorrect response on the pre-test. Neither offered an explanation on pre-test or post-test to justify their selections. Perhaps neither held a scientific nor specific non-scientific conception before instruction and that status had not changed after instruction. The teacher who stayed with option E showed little confidence in the justification provided on the post-test, which was that "the magnet will interact with other magnets" (4). This explanation was an improvement over "attracts all metals" (3), which was the pre-test explanation.

Figure 4: Task 4 assesses understanding of properties of magnets, specifically, that magnets have a N and S pole and strength of a magnet cannot be predicted by its size or shape.

4. Which of the following statements about bar, horseshoe, and round refrigerator magnets is most accurate?
   A. Large magnets are stronger than small magnets.
   B. Magnets have a N-pole and a S-pole.
   C. Horseshoe magnets are stronger than bar magnets which contain the same amount of material.
   D. Round magnets have only a N-pole or only a S-pole.
   E. A bar magnet will pick up more paper clips than a round refrigerator magnet.

Task 4 probes for understanding of magnets, including the concept that magnets have North-seeking and South-seeking poles and the fact that the strength of magnets cannot be predicted by their size or shape. Table 2 shows that 12 teachers selected the correct response on the pre-test and the remaining six teachers joined them on the post-test. Although only 7 of the 12 correct pre-test multiple-choice responses were supported with a satisfactory explanation, 14 of 18 correct responses were adequately supported on the post-test. Option C was the most popular distracter on the pre-test. Two of the four persons who selected it admitted to guessing and reported confidence levels of five. Another seemed to be responding to limited first-hand experience in explaining, "it just seemed stronger" (3), and the fourth teacher concluded, perhaps based on the appearance of the two shapes, "has 2X force on an object" (5), but expressed no confidence in the response. It is encouraging that option D was not selected by anyone. One of the 12 correct responders on the pre-test offered no explanation (3) and a second admitted to guessing (5). Two others simply indicated they had heard that statement before (2,2). The eight statements that adequately supported the scientific conception included "magnets have an N and an S pole regardless of size" (3) and "magnets have an N and an S pole regardless of shape" (3).

The increase in correct multiple-choice responses from 12 to 18, pre-test to post-test, for Task 4 was accompanied by a strong increase in confidence. Table 3 shows that only six teachers reported an initial confidence level of one or two on the pre-test, but all 18 participants did so

on the post-test. Interestingly, three persons who offered no attempt at an explanation and the one person who admitted to guessing all reported the top confidence level. Among the 14 satisfactory supporting explanations, two teachers wrote, "all magnets have N and S poles, regardless of size or shape" (1,1), and a third wrote, "to date all magnets have a north and south pole. Even if a north end is broken in two, the opposite end of the N becomes the south end" (1).

In Task 5 (Figure 5), participants can show understanding that the North pole of a bar magnet attracts not only the South pole of another magnet, but also attracts objects containing a ferromagnetic material, such as iron. Option D represents this response, and on the pre-test 11 of the 18 participants (61%) selected D. However, only three of the 11 gave an explanation that included two possible causes for the attraction. Of these three, only two made appropriate supporting comments, such as "opposites attract and iron attracts to magnets" (1). However, the third teacher wrote, "Metal will stick to the bar magnet and so will the S pole of another magnet" (1), which again supports the non-scientific conception that all metals are attracted to magnets, and the person was highly confident of that response. Of the other eight participants who chose the correct multiple-choice option on the pre-test, seven offered very brief statements that would be justification for selecting option B, such as "N attracts S" (3). The other respondent admitted to guessing. It seems likely that all of these teachers had observed a magnet attract objects not identified as magnets. However, the ferromagnetic material concept does not seem to be a functional component of most participants' conceptual framework for magnetism. Surprisingly, the confidence of these nine teachers ranged equally from top to bottom. Three were highly or somewhat confident, three were neutral in confidence, and three were somewhat lacking confidence or expressed no confidence.

Three participants (17%) chose B on the pre-test, which stated that attraction between opposite magnetic poles was "for sure" the reason that the identified magnet would be attracted to an unidentified object. One participant who chose B explained "opposites attract" (2), another "guessed" (5), and the third explained that "opposite poles attract as the electrons will align/bond" (2). Of the total of 18 participants, 14 (11 of whom selected D and 3 of whom selected B) indicated via multiple-choice selections on the pre-test that opposite magnetic poles attract. However, as noted, only a few explanations addressed the issue correctly and fully by invoking the type of evidence that might be gained through simple experiments. Finally, of the four participants who chose incorrect responses A, C, or E, only one (the one who selected option C) wrote an explanation, which was that "opposite poles repel each other" (4). The four participants who chose A, C, or E expressed neutral to low confidence.

The post-test multiple-choice results for Task 5 are both puzzling and disappointing. Nine teachers, compared to 11 on the pre-test, selected the correct answer (D). It is interesting that five of 11 teachers who selected the correct response on the pre-test migrated to option B on the post-test. Further, this is the only one of the five tasks for which movement from a correct multiple-choice response to an incorrect response occurred from the pre-test to the post-test. Migration to B also occurred from responses A and C. Could it be that participants were eager to complete the assessment tasks during the closing event of the institute and simply selected the option they considered to be the first plausible response? Note that all 18 participants selected either B or D on the post-test. Examination of the explanations provided by the nine persons who selected D showed that six participants clearly indicated that the N end of the object in the task could be an opposite pole, S, or the object could be made of a ferromagnetic material, such as iron. A representative supporting statement was: "The north pole of a magnet would attract the south pole of another magnet or any ferromagnetic material" (1). So, although the number of participants who selected the correct multiple-choice response dropped from 11 to nine from pre-test to post-test, there was an increase from two to six in the number of persons who chose the correct response and also provided a scientific explanation that identified both opposite poles and ferromagnetic material as plausible explanations for the attraction described in the test item. The small percentage of participants who both addressed the issue correctly and fully supported their response with evidence of the type that might be gained through simple experiments seems to reflect a deficiency in the instruction provided. One of the remaining three who selected the correct response on the post-test provided only the opposite pole explanation, another admitted to guessing, and the third did not offer any explanation. Of the nine teachers who selected option B, eight provided the opposite poles attract explanation and one provided no explanation. Looking at Table 3, confidence

reported in multiple-choice selections and supporting explanations moved from an average of 3.1 on the pre-test, a neutral level, to a high level average of 1.3 on the post-test.

For the five multiple-choice tasks combined, 49 of 90 responses (54.4%) were correct on the pre-test, and 75 of 90 responses (83.3%) were correct on the post-test. Further, multiple-choice responses were supported with satisfactory explanations for 25 of the 49 correct multiple-choice responses on the pre-test and for 56 of 75 correct multiple-choice responses on the post-test. Therefore, 25 of 90 multiple-choice responses (27.8%) on the pre-test and 56 of 90 (62.2%) on the post-test were both correct and cross-validated by explanations. These results were accompanied by sharply increased levels of reported confidence, as demonstrated by the finding that 34 of 90 ratings were 1 or 2 on pre-test compared to 84 of 90 on the post-test.

Figure 5: Task 5 assesses understanding that an unlike pole and ferromagnetic material are attracted to a magnetic pole.

5. Consider the diagram below

   The N-pole of a bar magnet is brought near end A of an object that looks very similar to the bar magnet in shape, size, and color. If end A of the object is attracted to the N-pole of the magnet, you could

   A. be sure that the object is another bar magnet and A is the N-pole.
   B. be sure that the object is another bar magnet and A is the S-pole.
   C. conclude that the object is either a bar magnet and A is the N-pole or the object is not a magnet but contains iron or a material that magnetically behaves like iron.
   D. conclude that the object is either a bar magnet and A is the S-pole or the object is not a magnet but contains iron or a material that magnetically behaves like iron.
   E. not make any of the conclusions in A. – D.

Conclusions and Implications

The results of this analysis indicate that the pre-institute level of the teachers' understanding of magnetism concepts had been overestimated, and, consequently, the extent and duration of the instruction needed was underestimated. Based on the results, it is concluded that the in-service elementary teachers in this sample had not previously received adequate content preparation to teach a rich unit on fundamental concepts of magnets and the behavior of magnets. Following completion of the short, non-traditional instructional intervention that was developed to be highly consistent with intentional conceptual change theory, the status of the group's conceptual understanding was much stronger but still in need of further improvement. Low confidence in multiple-choice responses and supporting explanations was frequently reported by teachers on the pre-test, even in instances when a correct response was selected. On the post-test, teacher confidence was much higher across all five tasks. Ideally, teachers would both provide evidence of strong science content preparation and be highly confident in their understanding of the content. The non-traditional instruction provided in this study seems to be associated with improvement in conceptual understanding and confidence in understanding. We conclude that the quantity of instruction provided should have been more extensive. The expectation that more extensive instruction with the same characteristics would be associated with evidence of better conceptual understanding is supported by a study of 178 pre-service elementary teachers who completed approximately 11 hours of magnetism instruction from *Physics by Inquiry*. Their pre-test performance on these same five tasks tended to be a little lower than the performance of the in-service teachers in the present study, but their post-test performance was essentially the same as found in the present study, except for Task 5. Only 41.6% of the pre-service group selected the correct response on pre-test for Task 5, but 84.8% did so on the post-test (Atwood & Christopher, 2007). In the present study, 61.1% selected the correct response on the pre-test, but only 50% did so on the post-test.

If teachers have not been adequately prepared to teach fundamental concepts about magnets and the behavior of magnets, there are important implications for both pre-service and in-service teacher education. First, it is likely that this topic is not being adequately addressed in pre-service teacher education programs, possibly because the topic is viewed as easier than other physical science topics. Alternatively, this might just be indicative of the more general problem of inadequate coursework in science for prospective elementary teachers (McDermott, 1991; McDermott et al., 2006; Trundle, Atwood & Christopher, 2002). In any case, a modest investment in appropriate instruction is associated with impressive gains in conceptual understanding (Atwood, Christopher, Combs & Roland,

2008). Additionally, the results of this study indicate that any assumption made by instructional supervisors or professional development providers that in-service elementary teachers are relatively well prepared to teach fundamental concepts about magnets and the behavior of magnets should be seriously questioned. It seems likely that the popular task of having children use magnets to test several objects in a classroom is not a highly productive activity in terms of concept development. Teachers leading these activities may not understand that all of the metallic objects interacting with a magnet almost certainly do so because they contain iron. (The odds of common metallic objects having nickel or cobalt in them are very small.) If teachers lack this knowledge, they are unlikely to help their students develop fundamental understanding by making sure a variety of non-ferromagnetic metals are identified and tested, followed by appropriate, sense-making discussions and explanations.

In addition, the results of this study suggest that teachers often lack experience determining where the magnetic effect of several magnets of varying shape and size is strongest (i.e., where the poles are located). By engaging in this process, teachers should determine that all magnets have two and only two magnetic poles and that like magnetic poles repel and unlike poles attract. Further, experience with large and small magnets of the same shape should be structured so it becomes clear that the strength of a magnet cannot be reliably predicted by size. Controlling size to the extent possible, while varying shape of magnets in appropriate investigations also seems to be needed for teachers to understand that the strength of a magnet cannot be reliably predicted by shape. Finally, more direct experiences and sense-making discussions about the effects that earth's magnetic field has on compass needles and other magnets seem to be needed for teachers. Clearly, these recommendations are not aligned with either traditional presentation mode instruction or with minimally-guided, non-traditional instruction. However, they are aligned with non-traditional instruction that is consistent with intentional conceptual change theory (Vosniadou, 1991, 2003, 2007).

Results of the present study could be used to help establish professional development priorities for in-service teachers and inform professional development plans that target magnetism and the behavior of magnets with instruction designed to promote conceptual change (Vosniadou, 1991, 2003, 2007). The results also could be used for formative purposes by higher education faculty who are committed to providing effective science programming for pre-service elementary teachers. The evaluation tasks fully described here could be used to determine whether other groups of pre-service or in-service teachers have essentially the same needs as were documented in the present study. Based on this study and literature cited earlier that documents the pervasiveness of the problem in the general population, we would predict this is not an isolated problem for either pre-service or in-service teachers.

Finally, we view one-on-one clinical interviews using props and probes as the most effective method of assessing the conceptual understanding of individuals (Trundle, Atwood & Christopher, 2002). However, in-service teachers are very wary of efforts to assess their content knowledge. This barrier, combined with a lack of sufficient time and other resources, make interviews of individual in-service teachers problematic and very difficult for instructional supervisors and other professional development providers to utilize. When supported by explanations and confidence ratings, multiple-choice tasks with popular non-scientific conceptions embedded in the distracter options, offer a viable alternative. The administration time is reasonable, and the data obtained can be very useful. Further, when a coding system is used to assure anonymity, teachers are comfortable and respond well to this mode of assessment.

References

American Association for the Advancement of Science. (2003). Benchmarks for science literacy. New York: Oxford University Press.

Arts, R. (2006). *A comparison of inservice elementary teachers' conceptions of selected standard-based force and motion concepts before and after instruction*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Atwood, R., Christopher. J. & McNall, R. (2006). *Middle school teachers' understanding of standards-based light concepts before and after instruction*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, April 3-7, San Francisco, CA.

Atwood, R. K. & Christopher, J. E. (2007). *Pre-service elementary teachers' understanding of standards-based magnetism concepts*. Paper presented at annual meeting of the Association for Science Teacher Education, January 3-6, Clearwater, FL.

Atwood, R.K. Christopher, J.E., & McNall, R.M. (2007). *Are inservice elementary teachers prepared to teach fundamental concepts of magnets and the behavior*

of magnets? Paper presented at annual meeting of the National Association for Research in Science Teaching, April 15-18, New Orleans, LA.

Atwood, R.K., Christopher, J.E., Combs, R.M., & Roland, E.A. (2008). *Preservice elementary teachers' understanding of magnetism concepts before and after instruction*. Paper presented at the annual meeting of the Association for Science Teacher Education, January 10-12, St. Louis, MO.

Bar, V. (1989). Children's views about the water cycle. *Science Education, 73*(4), 481-500.

Barman, C. R., & Griffiths, A. K. (1995). High school students' conceptions regarding food chains and food webs: a multinational study. *International Journal of Science Education, 17*(6), 775-782.

Barrow, L. (1990). Elementary science textbooks and potential magnet misconceptions. *School Science and Mathematics, 90*, 716-720.

Barrow, L. (2000). Do elementary science methods textbooks facilitate the understanding of magnet concepts? *Journal of Science Education and Technology, 9*, 199-205.

Baxter, J. (1989). Children's understanding of familiar astronomical events. *International Journal of Science Education, 11*, 502-513.

Beeth, M.E. (1998). Facilitating conceptual change learning: The need for teachers to support metacognition. *Journal of Science Teacher Education, 9*, 49-61.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honour of Robert Glasser* (pp. 361-392). Hillsdale, NJ: Lawrence Erlbaum.

Brody, M. J., & Koch, H. (1990). An assessment of 4th, 8th and 11th -grade students' knowledge related to marine science and natural resource issues. *Journal of Environmental Education, 14*(1), 1-14.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Champagne, A. B., Gunstone, R. F., & Klopfer, L. E. (1985). Effecting changes in cognitive structures among physics students. In L. West & A. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 163-188). Orlando, FL: Academic Press.

Constantinou, C., Raftopoulos, A., & Spanoudis, G. (2001). *Young children's construction of operational definitions in magnetism: The role of cognitive readiness and scaffolding the learning environment*. Paper presented at the Cognitive Science Conference.

Cosgrove, M., & Osborn, R. (1985). Lesson frameworks for changing children's ideas. In R. Osborn & F. P. Freyberg (Eds.), *Learning in science: the implications of children's science* (pp. 101-111). Portsmouth, NH: Heinemann.

Driver, R., Guesne, E., & Tiberghien, A. (Eds.). (1985). *Children's ideas in science*. Milton Keynes, England: Open University Press.

Duit, R. (1984). Learning the energy concept in school-empirical results from The Philippines and West Germany. *Physics Education, 19*(2), 59-66.

Duit, R. (2004). Bibliography: Students' and teachers' conceptions and science education database. Retrieved Available <http://ipn.uni-kiel.de/aktuell/stcse.html>.

Duit, R., & Treagust, D.F. (1995). Students' conceptions and constructivist teaching approaches. In B. J. Fraser & H. J. Walberg (Eds.), *Improving science education* (pp. 46-69). Chicago: The University of Chicago Press.

Finley, F. (1986). Evaluation instruction: The complementary use of clinical interviews. *Journal of Research in Science Teaching, 23*, 635-650.

Gallagher, J. J. (2000). Teaching for understanding and application of science knowledge. *School Science and Mathematics, 100*(6), 310-318.

Gardner, H., & Boix-Mansilla, V. (1994). Teaching for understanding in the disciplines-and beyond. *Teachers College Record, 96*(2), 198-218.

Hatano, G., & Inagaki, K. (2003). When is conceptual change intended?: A cognitive sociocultural view. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 407-427). Mahwah, NJ: Erlbaum.

Hennessy, M. G. (2003). Metacognitive aspects of students' reflective discourse: Implications for intentional conceptual change teaching and learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional Conceptual Change* (pp. 103-122). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *Physics Teacher, 30*, 141-158.

Hewson, M. G., & Hewson, P. W. (1983). Effect of instruction using students' prior knowledge and conceptual change strategies on science learning. *Journal of Research in Science Teaching, 20*, 731-743.

Hewson, P. W., & Hewson, M. G. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science, 13*, 1-13.

Hickey, R., & Schibeci, R. (1999). The attraction of magnetism. *Physics Education, 34*, 383-388.

Jackson, J., Dukerich, L., & Hestenes, D. (2008). Modeling instruction: An effective model for science education. *Science Educator, 17(1)*, 10-17.

Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experimental and inquiry- based teaching. *Educational Psychologist, 40*, 75-86.

Krall, R.M., Christopher, J.E., & Atwood, R.K. (2009). Comparison of Central Appalachian In-service Elementary and Middle School Teachers' Understanding of Selected Light and Force and Motion Concepts. *Science Educator*, 18(1), 1-16.

Mayer, R. (2004). Should there be a three-strike rule against pure discovery learning? *American Psychologist, 59*, 14-19.

McDermott, L. (1991). Milikan Lecture 1990: What we teach and what is learned: Closing the gap. *American Journal of Physics, 59*, 301-315.

McDermott, L. (1996). *Physics By Inquiry*. New York: John Wiley & Sons.

McDermott, L., Heron, P., Shaffer, P., & Stetzer, M. (2006). Improving the preparation of K-12 teachers through physics education research. *American Journal of Physics, 74*, 763-767.

National Research Council. (1996). National science education standards. Washington, D.C.: National Academy Press.

Niaz, M. (1995). Cognitive conflict as a teaching strategy in solving chemistry problems: A dialectic-constructivist perspective. *Journal of Research in Science Teaching, 32*, 959-970.

Nussbaum, J., & Novick, N. (1982). Alternative frameworks, conceptual conflict, and accommodation: Toward a principled teaching strategy. *Instructional Science, 11*, 183-200.

Oliver, J. S. (2007). Rural science education. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 345-369). Mahwah, New Jersey: Lawrence Erlbaum.

Osborn, R., & Cosgrove, M. (1983). Childrens' conceptions of the changes of state of water. *Journal of Research in Science Education, 20*(9), 825-838.

Osborn, R., & Wittrock, M. (1983). Learning science: A generative process. *Science Education, 67*(4), 489-508.

Prawat, R. S. (1989). Teaching for understanding: Three key attributes. *Teaching and Teacher Education, 5*(4), 315-328.

Schoon, K. J. (1992). Students' alternative conceptions of earth and space. *Journal of Geological Education, 40*, 209-214.

Sinatra, G. M., & Pintrich, P. R. (2003). The role of intentions in conceptual change learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional conceptual change* (pp. 1-18). Mahwah, NJ: Erlbaum.

Smith, L. E., Blakessie, T. D., & Anderson, C. W. (1993). Teaching strategies associated with conceptual change in science *Journal of Research in Science Teaching, 30*(2), 111-126.

Toleman, M. N. (1998). *Hands-on science activities for grades K-2*. West Nyack, NY: Parker Publishing Company.

Trundle, K.C., Atwood, R.K., & Christopher, J.E. (2002). Preservice elementary teachers' conceptions of moon phases. *Journal of Research in Science Teaching,* 39 (7), 633-658.

Vosniadou, S. (1991). Designing curricula for conceptual restructuring: Lessons from the study of knowledge acquisition in astronomy. *Journal of Curriculum Studies, 23*, 219-237.

Vosniadou, S. (2002). On the nature of naive physics. In M. Limon & L. Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice* (pp. 59-76). Dordrecht: Kluwer Academic Publishers.

Vosniadou, S. (2003). Exploring the relationship between conceptual change and intentional learning. In G. M. Sinatra & P. R. Pintrich (Eds.), *Intentional Conceptual Change* (pp. 377-406). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Vosniadou, S. (2007). The conceptual change approach and its re-framing In S. Vosniadou, A. Baltas & X. Vamvaskoussi (Eds.), *Reframing the conceptual change approach in learning and instruction* (pp. 1-15). Oxford, UK: Elsevier.

Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science, 18*, 123-183.

Weaver, G. C. (1998). Strategies in K-12 science instruction to promote conceptual change. *Science Education, 82*(4), 455-472.

Wildey, H., & Wallace, J. (1995). Understanding teaching or teaching for understanding: Alternative frameworks for science classroom. *Journal of Research in Science Teaching, 32*(2), 143-156.

Ronald K. Atwood is professor of science education , Department of Curriculum and Instruction, University of Kentucky, 335 Dickey Hall, Lexington, KY 40506. Correspondence concerning this article can be sent to <rkatwo00@uky.edu>.

John E. Christopher is associate professor of physics, Department of Physics and Astronomy, University of Kentucky, Lexington, KY 40506.

Rebecca K. Combs is research assistant, Department of Curriculum and Instruction, 335 Dickey Hall, Lexington, KY 40506

Elizabeth A. E. Roland is assistant professor of science education,405 C Lappin Hall, Department of Earth and Space Sciences, Morehead State University, Morehead, KY 40315

Information for Authors

The Science Educator, official journal of the National Science Education Leadership Association (NSELA), is a refereed journal which seeks manuscripts dealing with topics and issues of interest to professionals involved in science education leadership across a variety of roles, institutions, and agencies. The journal is published bi-annually with provisions for quarterly publication. The journal serves as a vehicle for the exchange of information on current theory, research, and classroom applications. Articles by both practitioners and researchers are encouraged. Generally, articles are sought that aid practitioners in making more informed decisions about current or future operations, and which improve science teaching and learning, at all levels, in our nation's schools. Of particular interest are topics that deal with the following issues: curriculum and instruction, science education reform, science education leadership, preservice and inservice teacher enhancement/professional development, teaching and learning, needs assessments/research and evaluation, and polices and practices.

To have a manuscript considered for publication, please review the guidelines below:

1. All manuscripts should be sent electronically to <brenda.wohnowski@unt.edu>; manuscripts should not exceed 5000 words. Samples of previously published articles in the Science Educator can be viewed on the NSELA homepage <www.nsela.org>.
2. Manuscripts should be typed double-spaced on standard 8.5 x 11 inch paper. Figures, tables and pictures must be submitted in camera ready condition. Tables should be typed on separate pages at the end of the manuscript.
3. Manuscripts must be concisely written. It is important to maintain coherence of thought.
4. The first page of the manuscript should contain the article title, author's name, affiliation, and address to which correspondence and proofs should be sent. The author should not place his/her name on the other manuscript pages so that anonymity can be maintained during the review process.
5. References cited in the manuscript should be listed at the end of the manuscript. Provide accurate and complete information. For questions of format follow the Publication Manual of the American Psychological Association (Fifth Edition) and Webster's College Dictionary.
6. Authors who are not members of the National Science Leadership Association (NSELA) will be assessed a publication fee of $50.00 per article. Non members may avoid this charge by joining NSELA prior to their manuscripts being published. The institutional charge is $75, $30 individual journal-only. Membership dues are payable at the beginning of each calendar year. Information and application for membership can be obtained from the NSELA Office Manager Judy Hamilton <judyhamilton51@gmail.com> or Membership Chair <beth@seven-oaks.net>.
7. Receipt of manuscripts will be acknowledged. All submitted manuscripts will be reviewed by at least three members of the Editorial Board. Authors will be notified by the Editor regarding the recommendation on publication of the manuscript. Transfer of the author's copyright to the National Science Education Leadership Association is a condition for publication.

ExploreLearning

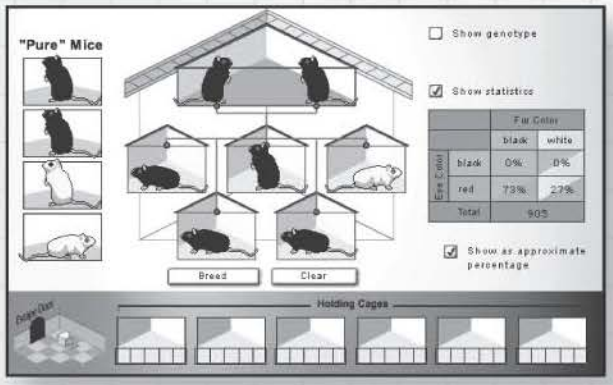
EXPERIENCE THE **Ah-hah Moment** WITH GIZMOS

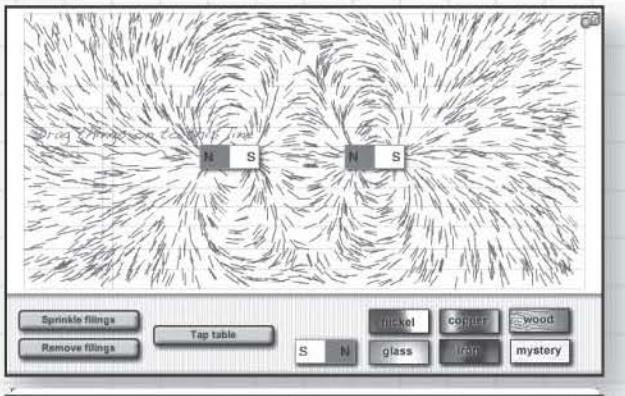
Unleash the power of inquiry-based learning using the award-winning library of Science Gizmos at ExploreLearning.com. Designed for grades 3-12, Gizmos help you take advantage of research-proven instructional strategies that let students of all ability levels develop conceptual understanding. With over 450 Gizmos at your disposal, you can supplement and enhance your instruction with powerful interactive visualizations of science concepts. Students can manipulate key variables, generate and test hypotheses, and engage in extensive "what-if" experimentation.

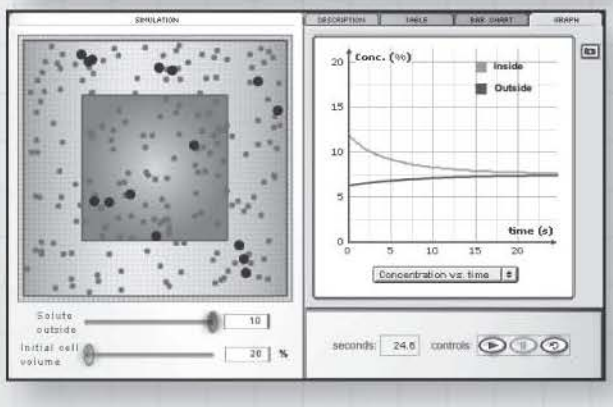
Key features include:

- *Inquiry-based lessons that are ready to use or customizable*
- *Correlations to state and national standards and over 200 textbooks*
- *Formative assessment with instant feedback for students and real-time reports for teachers*
- *Easy-to-use product so time is spent teaching and learning science, not the technology*
- *Teacher guides with answer keys to support elementary teachers in Gizmos for grades 3-5.*

Visit www.explorelearning.com for more information or to take a FREE 30-day trial of Gizmos







P.O. Box 2185 • Charlottesville, VA 22902 sales@explorelearning.com 866.882.4141 877-829-3039