The Delaware LSCI: A Systemic Partnership to improve the Teaching and Learning of Science for All Children. (Award ID 9618984)

FINAL REPORT TO THE NATIONAL SCIENCE FOUNDATION

Rachel E. Wood and John W. Collette

Objective of the Delaware LSCI

The objective of the Delaware Local Systemic Change Initiative is to build capacity at the school, district and state level to ensure that all children in grades K-8 have the opportunity to obtain the skills and knowledge needed to meet the expectations of the Delaware Science Standards. The LSCI is the responsibility of the Delaware Science Coalition, a coalition of 15 Delaware school districts in partnership with the Delaware Department of Education (DOE) and the Delaware Foundation for Science and Mathematics Education (DFSME), a business supported non-profit organization committed to support implementation of the Delaware Science and Mathematics Standards

Program Background Prior the LSCI Grant

The Science Coalition program began in 1995 as a pilot initiative program in seven school districts (six rural and one suburban) shortly after the adoption of the Delaware Science Standards in June 1995 (Ref 1). The K-6 pilot was designed by a team of educators and community representatives after attending a Leadership Training Institute in Elementary Science Education Reform offered by the National Science Resource Center (NSRC) in Washington DC and is based on the NSRC model for systemic reform (Ref 2).

The pilot was designed to determine how effective a modular (i.e. kit-based) inquiry-centered curriculum would be in enabling Delaware teachers to meet the Delaware standards. The plan included providing each teacher with the classroom materials and intensive professional development in the science unit.

Funding for this pilot came from the participating districts, DFSME (representing the local business community) and a special grant from the legislature. In the first year, the Coalition:

- Implemented one curriculum unit in each K-6 grade aligned specifically with the grade level Performance Indicators (Ref 1).
- Provided professional development to 356 K-6 (ca. 50 teachers/grade) in the appropriate science unit.
- Engaged about 7000 students in hands-on inquiry-centered science instruction.
- Established a temporary science resource center to store materials and engaged volunteers to refurbish the kits as needed.

The program was an instant success with teachers, students, parents and administrators. The students responded enthusiastically to the hands-on science. Teachers found the classroom materials greatly facilitated instruction and rated the professional development as some of the best they had ever had. Some parents became active advocates as a result of their children's enthusiasm. Administrators agreed that this approach was effective in preparing teachers to meet the Standards.

This led to strong district and school support for continuing and expanding the program. Two more school districts (one rural and one suburban/ urban) joined the Coalition in 1996 significantly expanding the size of the program and the demands on the small staff to provide increasing numbers of professional development courses. While supplemental funding was obtained from the legislature and the business community for the second year, it was clear that external funding and additional staff would be required to meet the need for more materials and more professional development. The Science Coalition submitted a proposal to NSF for a Local Systemic Change Initiative grant that was funded for the 1997 school year.

The LSCI Proposal

The specific goals and activities of the LSCI are as follows:

1. Establish and continuously improve an inquiry-based K-8 science curriculum that provides every student with the opportunity to develop the skills and knowledge needed to meet the Delaware Science Standards
2. Provide in-depth professional development to all K-8 teachers of science so they can become proficient instructors of inquiry-based science.
3. Establish materials support system so that teachers have the classroom materials needed for all children to participate in inquiry-centered instruction.
4. Develop a comprehensive assessment system that informs teachers and administrators of student learning.
5. Establish and nurture a network of teacher leaders who will lead and sustain science education reform at the school and district level.
6. Build and sustain a process to ensure that school, district and state policies are supportive of systemic reform and that community support is effectively focused on common programmatic objectives.

ACTIVITIES

Previous reports to NSF have detailed the progress during the first four years of the LSCI. Through August 2001, the Delaware Science Coalition:

- Expanded to include 15 of the 16 districts in Delaware serving children in grades K-8. One district did not sign the Memorandum of Agreement for unique local reasons. However, the district adopted the same kit-based curriculum, purchased and refurbished their kits. District personnel have worked closely with Coalition Leadership and arranged to have their teachers participate in the professional development offered by the Coalition.
- Established a leadership structure including Science Specialists in every school district and Lead Teachers in every school to lead and support implementation at the district and school level respectively.
- Began to develop summative assessments for all K-5 units as a key component of a comprehensive assessment system that could provide timely diagnostic information to teachers about student learning.

- Expanded the LSC to middle school in 1998 with an agreement by all school districts to establish an inquiry-centered Grade 6-8 curriculum and to provide teacher professional development to all middle school teachers of science.
- Developed a strategic plan to sustain the Science Coalition after NSF funding ended. This was the most critical issue facing the Coalition at the beginning of the fourth year of the program. All participants agreed that the Coalition program was making progress in establishing a comprehensive Standards-based curriculum, improving teacher quality and student learning. Recognizing that much more needed to be done, all parties agreed on a new five-year plan and signed a new Memorandum of Agreement that not only reaffirmed the commitment to complete the K-8 curriculum but also added a commitment to begin expanding the program to high school.

This report discusses progress since September 2001 as well as the overall impact and contributions of the LSC. State testing of science began in fall 2000 through the Delaware Student Testing Program (DSTP). This provided direct measures of the effectiveness of the LSCI in raising student achievement in science and helped identify areas for further work. The first DSTP results became available in Spring 2002.

A. Restructuring The Coalition for Sustainability

The strategic plan developed during the fourth year was implemented over the next year. The no cost extension given to the LSCI by NSF provided additional time to transition to a sustainable system.

As a first step in implementing the new plan, the LSCI leadership was restructured to provide increased attention to the unique needs of the elementary and middle schools and to make more effective use of the varied talents and expertise of the Science Specialists. The Core Leadership team was expanded to include three Science Specialists as Teachers on Special Assignment focused on grades K-5, including one with a Special Education background; and three Science Specialists focused on middle school curriculum and assessment. A Math Specialist was engaged to build links between middle school mathematics and science curricula and teachers.

The responsibilities of the district Science Specialists were changed so that could devote more time to support teachers in the classroom and to strengthening teacher professional development. The Science Resource Center took over full responsibility for logistic support of science kits, a task that previously been handled by the Specialists.

B. Improving the Middle School Curriculum

All units of the 6-8 curriculum units have been completed and kit usage data show teachers throughout the state are using them in the classroom. We recognize, however, that this curriculum is still a work in progress. Many of the units were developed before there was an adequate supply of commercial materials. All undergo continued review and revision as more teachers gain experience with them in their classrooms, identify gaps that are not covered or find units that cover the standards more effectively. The summative assessments are also driving curriculum improvement by identifying curricular gaps and student misconceptions. We expect that this process will continue for some time. The development process is being aided substantially by the active participation of faculty members from the University of Delaware and Delaware State University and experts from DNREC.

C. Building Assessments and Assessment Literacy

The LSCI continues to focus on the development of summative assessment tailored to each inquiry-based curriculum unit as one component of a comprehensive assessment system. These diagnostic assessments complement the state science tests providing a low stakes test that can provide teachers with timely feedback about what students are or are not learning. Each assessment is developed by a small group of grade-level teachers working with Coalition leadership and advised by Dr Mary Ellen Harmon of Boston University and then piloted by Lead Teachers. The assessments have become an important component of teacher professional development providing a very effective way to engage teachers in discussions of their own students' work and of ways to improve instruction.

The elementary Lead teachers have completed development and testing of summative assessments for every K-5 curricular unit. They are now in the process of building a core of teachers in every school who understand and know how to administer the assessments, score student work using double-digit rubrics, draw valid inferences from the data and use the assessments to guide instructional improvement.

The elementary lead teachers have begun to work on formative assessments for each curricular unit. The formative assessments are developed around pivotal lessons that best reflect the "big ideas" in the unit. They are designed to help teachers understand the types of questions they need to ask to determine to what extent their students understand these ideas and to inform their instructional strategies. Summative assessments have also been developed for one middle school unit at each grade. These are being disseminated through the middle school Lead Teachers and being used in conjunction with the middle school teacher professional development program. Summative assessments for all other middle school units will be developed over the next two years. The assessments are available at **www.scienceassessment.org**.

All Lead teachers have taken professional development in the analyses and interpretation of the DSTP student achievement results so they can support their colleagues and peers in the use of statewide data.

D. Strengthening and Expanding Professional Development

Elementary Professional Development

Most professional development courses are provided for K-5 teachers with about 1300 teachers participating in courses in the 2003/4 school year. We had anticipated that the number of teachers taking such courses would decrease over time. To our surprise participation in professional development has stayed at a level of 1300 to 1400 teachers/year in the period from 2001 through June 2004. (Appendix, Figure A1). This means that about 40% of the K-5 teachers are taking professional development annually and that, through June 2004, more than 12,256 K-5 teachers have taken about 250,000 hours of professional development in science.

A number of factors contribute to this high level. The most important is the continued high rate of teacher turnover (20-30% annually) resulting from teachers leaving the system or moving to different grades. District administrators know that the capacity of their teacher workforce to teach science will degrade rapidly if new teachers do not participate in the professional development. A second factor has been the effort by Coalition districts to include elementary special education teachers in the science training; this represents a new group of teachers that had not been explicitly included in the early part of the program. Finally the state has recently established new licensure requirements so all teachers must now take regular professional development each year.

The Science Specialists who lead the elementary program have systematically improved the professional development over the last three years. This was driven by our own evaluations and those of our external evaluators that indicated that they the professional development was weak in the area of deepening teacher science content knowledge; and that there was considerable variability in the effectiveness of different instructors.

This quality improvement program established more definitive guidance to instructors about the learning expectations for each unit and clarified their responsibilities for ensuring that the key science concepts were covered and understood. It included regular observations of instructors and systematically helping or replacing the less effective ones. Alternatives to the professional development format of after-school training were added to meet different teachers needs and to allow teachers more time for reflection and for modeling effective pedagogy.

We believe that these initiatives have contributed to the improvements in student achievement at grade 4 and grade 6 since 2000. To further strengthen teacher content knowledge, a set of cluster courses have been developed that complement the kit training with courses on assessment, science content and teacher leadership; this cluster will also meet teacher's licensure requirements.

Middle School Professional Development.

The demand for professional development from middle school teachers increased significantly in the last three years rising from 127 in 2001 to 445 in 2003, representing a total of ca. 50,000hours of professional development since the program began in 1998 (Fig A1). The recent increase is due to the availability of more 6-8 curriculum units, teacher turnover and a significant increase in the number of secondary special education teachers who have recognized the need to improve their content knowledge to adequately prepare their students to meet the science standards.

Middle school courses consist of 37-42 hours of professional development or an equivalent graduate level course offered by one of the universities in the summer. The courses are typically co-taught by a content expert working with a classroom teacher. The experts are drawn from university faculty (including the Colleges of Arts and Science, Marine Studies, Agriculture and Education) or from a specialized government agency such as Delaware Natural Resources and Environmental Control Agency(DNREC). Teachers can use the courses as credit to a masters degree program offered by the University of Delaware. The Mathematics and Science Education Resource Center (MSERC) at the University of Delaware also provides follow up support in the form of classroom visits or e-mail and phone consultations for any courses they provide.

Because the demand for these courses continues to exceed the current supply, the Coalition is working with these institutions to expand the number of courses available during the academic year**.**

FINDINGS

A. Summary

The specific findings outlined below document that the Science Coalition LSCI has met its goal of contributing to systemic improvement of science education in Delaware. Specifically:

- A challenging standards-based K-8 science curriculum has been established, adopted and is being implemented in all Delaware public schools.
- A professional development system has been established that provides all Delaware K-8 teachers with access to the content, pedagogical and assessment knowledge and skills required to teach inquiry-based science.
- Student achievement in science has been raised for all students.
- The achievement gap between white and minority students has been reduced and in some grade 4 schools has even been eliminated.

The bottom line is that, as a result of the LSCI, the majority of Delaware students are getting more science than ever before, the science is based on a more challenging and more coherent curriculum that is being delivered by better prepared teachers and is leading to significant improvements in student achievement.

B. Widespread Adoption and Implementation of the Standards-based Curriculum

The Standards-based K-8 curriculum developed by the Coalition has been adopted by all participating school districts and by five Charter schools. This curriculum can be viewed at www.sciencede.org.

The implications of this rather simple statement should not be underestimated. It represents a major change from the situation before the Coalition and before the LSCI. Not only do all students get a full science curriculum in Grades K-5 where little science was taught before, the entire K-8 curriculum is much more challenging than it was in 1995. Topics that were typically taught in Grade 7 (Ecosystems, Mixtures and Solutions) are now part of the Grade 5 curriculum. Electricity was often not covered until high school; today students study Electric Circuits in grade 4 and have to demonstrate their understanding of the subject on the DSTP. The curriculum units are carefully sequenced so that more advanced ones build on prior knowledge leading to continual growth in a student's knowledge base and reducing the need for repetition across grades.

One of the more difficult aspects of effecting these structural changes has been the challenge of convincing middle school teachers that they have to give up "their" topic as it is now being covered at an earlier grade. This same process of change is currently facing high school teachers who find that subjects traditionally taught in grade 9 are now part of the middle school curriculum.

The best measures of curriculum implementation are the number of teachers who have taken professional development in the curriculum units (Appendix, Fig A1) and the number of kits distributed by the Science Resource Center. In the 2003/4 school year, the

SRC distributed 5300 kits to K-5 teachers and 684 kits to 6-8 teachers. The SRC reports that only a small percentage of these return unopened.

Surveys carried out as part of the Core Evaluation show that 80% of the elementary teachers report teaching 3-5 science lessons/week with lessons of 40-45 minutes (Ref 3). This compares to a 1995 survey that found that science in the early elementary grades was limited to about 35 minutes/week

These data are very positive and undoubtedly reflect the extent to which the science curriculum has been broadly adopted across the state. It is unrealistic, however, to believe that the situation is without problems, but the concerns are more likely to deal with how well the curriculum is being implemented rather than what, if any, science is being taught. An example of this was obtained through a recent study of student achievement in Grade 4 schools in which teachers and students in three elementary schools were interviewed in detail. This found wide variations among individual teachers in each school both on the number of science lessons taught each week and in the length of the lessons (Ref 4). Another uncertainty is the extent to which teachers complete the unit without dropping pivotal lessons for lack of time or a failure to recognize the importance of the lesson. As discussed in more detail below, the performance of individual schools vary widely across the state. These differences are very likely related to how effectively the curriculum is being implemented in the classroom.

C. Significant Changes in Instructional Practices

The Core evaluation confirms that the LSC has made substantial changes in the quality of classroom instruction in terms of using the designated materials and that *"teachers are no longer teaching science from a textbook and are using more hands-on investigation with the students"(*Ref. 3*)*. The evaluation found that 90% of the survey respondents have very positive attitudes and beliefs about science teaching and learning and 87% felt well prepared to teach science. Further, 90% of the classroom teachers interviewed reported that they had changed classroom practice because of the project. When asked about successful strategies and activities of the Coalition, one CSS remarked that it was *"removing the fear of elementary teachers to do science."*(Ref 3)

The results from student surveys given with the DSTP provide supporting evidence of the impact of the science curriculum on students (Zhang, Ref. 5). Zhang reports that:

- 69% of students in grade 4 and 6 reported using scientific equipment (balances, microscopes, thermometers, hand lenses) daily or weekly in 2003. This is an increase from 55% in 2000.
- 57% of grade 4 and 70% of grade 6 students had to write explanations in journals or lab reports in 2003. This has increased from 48% and 59% respectively in 2000.

Zhang found that both of these variables showed a positive relationship with student performance in science in Grades 6 and 8.

D. NSF funding has been significantly leveraged in meeting the goals of the LSC to improve the teaching and learning of all students.

A total of $23.7m has been spent in the Science Coalition K-8 program since its inception in Sept 1995. Of this total, NSF contributed 25.5%, the state contributed 29,8%, the participating districts contributed 41% and DFSME contributed 3.5%. The totals do not include other significant funding for science that has been contributed to individual teachers or schools by MBNA, DuPont or other local companies independent of DFSME.

This shows that the NSF funding of $6,050.000 has been leveraged a factor of 4 in meeting the stated objectives of the LSC for systemic improvement in science education in Delaware.

E. Student Achievement has Increased for All Students.
Background

The most important findings from this work are the successes in raising student achievement of students in Grades 4 and 6. These increases are clearly related to the LSCI. There is no other comprehensive program active in Delaware specifically aimed at improving student achievement in science and there is certainly no other one that reaches so many schools and students across the state.

We have used student achievement on the state tests given through the DSTP as the main measure of the program's impact (Ref.1). These tests are given at the beginning of Grade 4 (measuring performance in grades K-3), the beginning of Grade 6 (measuring performance in grades 4 and 5) and the end of Grade 8 (measuring performance in grade 6,7 and 8).

Student performance is divided into five Performance Levels based on student scale score: Distinguished (PL 5), Exceeding the Standard (PL 4), Meeting the Standard (PL 3), Below the Standard (PL 2) and Well Below the Standard (PL1). (See Appendix B for more details). We have used the sum of PL3, 4 and 5 (or the percent students meeting/exceeding the standard) as the main measure for discussing the student achievement results. These PLs are linked to scale scores and can provide much more information about the breadth of the distribution of scores than the average raw or scale scores do. And they have a very real meaning to students, educators and the public.

We have used the change of student achievement over time as the best measure of improvement and of the impact of the LSCI. While we have been able to relate improvements in grade 8 student achievement to specific teacher professional development, we were not able to demonstrate such linkages in grades 4 and 6. There are several reasons for this:

1. The state science test did not become available until 2000, five years after the pilot program started. By this time, a large percentage of teachers had taken professional development leaving us with no base case from which to measure.
2. We had not anticipated the extent and the impact of student movement among schools. One study of three grade 4 schools found that more than 60% of grade 4 students had attended other schools than the one in which they took the Grade 4 (Ref 4); most Grade 6 students in the state are tested in a school different from the one they attended for Grades 4 and 5. Data systems were not readily available at that time to link these students retroactively with the teachers who taught them science.

We have examined in some detail the impact of the LSCI on the achievement gap between white students (W) and African Americans (AA). Many studies have shown that raising the performance of non-Asian minorities is critical to improving overall student performance in any school system. The W-AA gap has been used because African

Americans make up about one-third of the student population in Delaware and are well represented in most schools. Hispanics make up about 6% of the student population but many schools do not have enough Hispanic students to provide statistically valid student performance results.

There is no statistically significant difference in the achievement of males and females at either Grade 4 or Grade 6. Males outperform females slightly at Grade 8, but the differences are much smaller than those related to race.

Student achievement has increased for all students in Grades 4, 6 and 8 since student testing began in 2000. (Appendix, Fig A2).

Figure A2 shows a steady increase in the percent students meeting/exceeding the standard at each grade for the period 2000 to 2003:

- At grade 4, the percent meeting the standard increased from 85.0% in 2000 to 89.7% in 2002, but dropped to 88.7% in 2003 for a net increase of 3.7% indicating almost nine out of every ten Grade 4 students are succeeding in science. This is an exceptional accomplishment given the status of elementary science before the LSCI. It is unfortunate that we were not able to get measurements of student achievement earlier in the project. The drop in 2003 may reflect increased attention to Grade 3 mathematics (with less on science) to meet the Grade 3 mathematics accountability requirements.
- At Grade 6 student achievement increased from 63.4% in 2000 to 74.4% by 2003, an increase of 11%.
- At Grade 8, only 42% of the students met the standard in 2000. This increased marginally to 44% by 2002, and to 49% in 2003. These low percentages undoubtedly reflect the fact that the Grade 6-8 curriculum has just been completed and that many middle school teachers have not taken all the necessary professional development in the new units. The encouraging increase in student achievement in 2003 is mainly the result of a specific professional development program that is described in more detail below.

Student achievement for special education students increased at a faster rate than that of regular students from 2000 to 2003 resulting in a decrease in the achievement gap between special education and regular education students in grade 4 and 6. (Appendix, Figures A3-5).

The LSCI focused on the needs of special education students from the project inception. These students make up 10-20% of the student population in Delaware elementary schools and typically score lower than regular education students. In the past, many special education students have been routinely excluded from elementary science classes. The LSCI established a principle that all special educations students would be given science instruction and would be included in regular science classes.

The impact of this approach was confirmed during the first two years of the project by an independent evaluator who reported that science was the only subject she observed in which she could not differentiate special educations students from regular students in typical classroom situations (Ref. 6). The positive response of special educations students to the hands-on learning approach played an important role in engaging parental support in the early stages of the program.

It is therefore very rewarding to document that the achievement of special education students has increased more rapidly than that of regular education students in Grades 4 and 6 in the 2000- 2003 period (Figure A3) substantially reducing this performance gap. There has been little or no progress at Grade 8, but an active program is now in place to provide extra professional development to middle school special education teachers.

The improvements achieved with special education students between 2000 and 2002 are greater in science than in other disciplines. Figures A4 and A5 compare the changes in Grade 4 and 6 science scores (respectively) with those in mathematics, reading and social studies. The absolutes scores among these disciplines cannot be compared as they involve different tests and are given at different times. However, the relative changes between 2000 and 2002 show that, at Grade 4, science and reading raised the performance of special education students relative to regular students at a faster rate than was done in mathematics or social studies. At grade 6, the performance of special education students relative to regular students increased at a faster rate in science than any other discipline.

The comparisons between science and social studies are particularly relevant as the same teacher in the elementary grades often teaches these subjects and the tests are given at the same time. The major differences between these two subjects should therefore relate primarily to the common science curriculum in the state and the increased level of professional development that science teachers must take.

The achievement gap between white and minority students decreased in Grade 4 and Grade 6 from 2000-2003. (Appendix, Figures A6-8).

- At Grade 4, the achievement gap between white and African American students dropped from 22.6%(2000) to 16.4% (2003). The achievement gap between white and Hispanic students dropped from 18.1% to 13.0% (Fig A6). These encouraging developments are an important demonstration of the impact of the LSCI. The individual school data discussed below are even more positive.
- At Grade 6, the gap between white and African American students dropped from 37% (2000) to 30% (2003) while the white-Hispanic gap dropped from 29.3% to 24.6% (Fig A7).
- At Grade 8, while student achievement increased for all groups, the white-African American achievement gap actually increased from 34.9% to 36.8%, while the white-Hispanic gap decreased from 34.9% to 31.8% (Fig. A8). These levels are all clearly too high, and undoubtedly reflect the current implementation stage of the middle school program. As discussed below, encouraging progress has been made in demonstrating how to reduce the gap further.

The achievement gap between white and African American students has been significantly reduced and even eliminated in a number of Grade 4 schools,

Regression analysis of the 2002 grade 4 data showed that student achievement was strongly correlated with race (R^2=0.57) and with the percent of low-income students (R^2=0.57). (Low-income students are students on free or reduced-price lunch). Such correlations have been documented in many other studies. However, a detailed analysis of individual school results revealed a more complex, more informative and more encouraging picture showing that, while race and income may be significant variables, they do not explain the range of variations among Delaware schools.

The 84 Grade 4 schools can be segmented into three distinct groups based on their racial diversity or lack of it. Table 1 shows that 76% of the students are in 56 schools with racially diverse student populations; 17% of the students are in 17 high poverty, high minority schools; and 7% of the students are in 11 low poverty, low minority schools. No achievement gap data are available for schools classified as low minority and high minority as they do not have enough students of one race to provide a valid measure of their average achievement.

Student achievement varied considerably among these three segments:

- Based on average school score, all the low poverty, low minority schools performed at the top and most, but not all, of the high poverty, high minority schools performed at the bottom. While this is consistent with the regression analyses, there are exceptions; one of the high poverty schools with 92.6% of its students meeting the standard clearly outperformed some of the low poverty schools
- The racially diverse schools performed across the whole range. They outperformed some of the low minority schools at the top and the outperformed all high minority schools at the bottom.

TABLE 1
Student Achievement(2002) In Different Grade 4 Schools (1)

SCHOOLS	State	Diverse Schools(2)	High Poverty High Minority (3)	Low Poverty Low Minority(3)
Number of Schools	84	56	11	17
Number of Students (% total)	8509	6177(76)	573 (7.0)	1407 (17)
STUDENT CHARACTERISTICS				
% Low Income(5)	44.1	47	86	21.9
% White	56.4	55	4.2	82.5
% African American	33.9	34.2	81.6	9
% Hispanic	6.8	7.4	13.4	4.7
SCHOOL PERFORMANCE				
% Students meeting				
Mean (stnddev)	89.6	90.6(4.8)	74(11.5)	97.2 (2.6)
Minimum	54.9	78.2	54.9	89.1
Maximum	100	98	92.6	100
% White Students Meeting	95.4	95.1		97.8
% AA Students Meeting	80.7	84.1	73	
Achievement Gap (Wh-AA)	14.7	11(4)		

NOTES
1. Based on 2002 science DSTP results.
2. These schools report achievement data on both races.
3.These schools report achievement results only for one race
4. Includes 14 schools with gap of <5.
5. Students on free or reduced price lunch.

The racially diverse schools provide a more rigorous comparison of the performance of the two races in the same environment. The regression correlation coefficients between student achievement and race (R^2 =0.25) and low income (R^2=0.19) were much lower for these schools than those when all 84 schools were included, reflecting the absence of data from the extremes. The average student achievement in these schools is higher than the state average and the achievement gap is lower. The achievement gap for these schools showed no correlation with either race or percent low-income students.

Table 2 compares the ten highest and the ten lowest performing diverse schools both in terms of the average student achievement and in terms of the achievement gap. The comparison by achievement shows that the lowest achieving schools have more low-income students (53% vs. 39.4%) and a higher percentage of minorities (52.1% vs. 34.6%). While student achievement for both white and African American students is lower in these schools than in the top ten schools, more than 90% of the white students still meet the standard. The average achievement of African American students drops much more than that of white students leading to an increase in the achievement gap.

TABLE 2

Comparison of Ten Highest and Ten Lowest Performing Grade 4 Schools(1)

	By Achievement			**By Achievement Gap**	
SCHOOLS	Highest Achieving	Lowest Achieving		Lowest Gap	Highest Gap
Urban/Rural Ratio	4/6	9/1		3/7	9/1
Number of Students	990	1525		1041	1105
STUDENT CHARACTERISTICS					
% Low Income(2)	39.4	53		47.6	46.5
% White	61,8	46.1		51.2	52
% African American	29	42		37.8	32.4
% Hispanic	5.6	10.1		8,2	11.6
SCHOOL PERFORMANCE					
% Students meeting	96.8	82.9		93.2	87.1
% white meeting	98.5	91.8		94.5	96.5
% AA meeting	92.8	74.6		93	71.8
Achievement Gap	5.8	17.2		1.5(3)	24.7

NOTES
1. Based on 2002 science DSTP results.
2.Students on free or reduced-price lunch.
3.Includes four schools with a Gap< 0.

The comparison of schools by the size of the achievement gap provides a somewhat different perspective. Student demographics are virtually the same in both the schools with the lowest and highest achievement gap. In the low gap schools (average gap of 1.5), more than 90% of the students meet the standards and there is essentially no difference between the average achievement of white and African American students. In four of these schools, African American students had higher average achievement than did white students resulting in a negative gap. In the high gap schools, the achievement of white students is marginally higher than that of the low gap schools (96.5% vs. 94.5%), but the achievement of African American students is much lower.

The results on the low gap schools clearly show that the LSCI program can reduce racial disparity so that all children can succeed. Further studies are now in progress to understand the factors that contribute to lower performance in some of these diverse schools. The low gap, high achieving schools are split evenly between rural and urban areas whereas nine of the ten lower performing schools are in urban environments. The lower scores in these schools may reflect differences in the extent and the integrity of implementation in the classroom as observed in mathematics by Briars and Resnick (Ref 7).

Encouraging leads have been found for reducing the achievement gap increase in Grade 6 relative to Grade 4.

Comparison of student achievement of the same cohort of students over time can provide a longitudinal measure of the impact of the LSCI and has been recommended as one of the best ways to identify programs that really work (Dougherty, Ref 8). This type of measurement has to be used with caution in the current situation as the science scores on the Grades 4,6,8 and 11 tests are not linked by a common scale and are therefore not directly comparable (Zhang, Ref. 9). Relative comparisons of longitudinal changes in student achievement and changes in the achievement gap between whites and minority students can however be compared and be informative.

In this context, student achievement in science shows an <u>apparent</u> drop in performance from Grade 4 to Grade 6 and from Grade 6 to Grade 8 for the same cohort of students while the achievement gap between white and African American students increases significantly in grades 6 and 8 relative to that in grade 4. From Grade 4 in 2000 to Grade 6 in 2002, average student achievement dropped from 85% to 73.6%. While performance of both races was lower in grade 6, minorities dropped a greater degree, so the achievement gap increased from 22.6%(2001) to 29.6%(2003. This is a serious concern as it suggests that the science program is not working as well with minorities in grades 4 and 5 as it was with white students.

Disaggregation of the 2002 data for all 38 grade 6 schools showed that the grade 6 schools fell into a very different pattern than the Grade 4 schools. The schools are larger, and 34 of the 38 schools have a diverse student population for which achievement data are available for both races. Student achievement by school varied from 40% of the students meeting to 100% and the achievement gap ranged from 7% to 52%. The majority of the lower performing schools are in the large urban districts. The evidence below strongly suggests that one factor is that the Science Specialist in these districts has more limited impact on individual teachers because of the larger number of different schools and teachers she or he has to work with.

All but one of the fourteen school districts showed a drop in nominal student achievement between grade 4 and grade 6 and an increase in the achievement gap. The one exception, Woodbridge, a rural district with one K-8 school, had a dramatically different result. Nominal student achievement for Woodbridge students actually increased from Grade 4(2000) to Grade 6(2002) and the achievement gap dropped (Table 3). We have examined several factors that could account for these differences:

a). Woodbridge had a relatively stable student population whereas the state student population increased by about 10% from 2000 to 2002; other data show that these incoming students typically perform at a lower level than do the current students. However, as shown in Table 3, an analysis of matched student results for the state between 2000 and 2002 showed a 9.7% drop in student achievement and a 6.9% increase in the achievement gap comparable to that for the whole state (Zhang, Ref.9). This compares to Woodbridge results of a 3.2% increase in student achievement and a 10.3% drop in achievement gap.

b). A second factor that could contribute to the differences is that most students in the state change schools between Grade 4 and Grade 6 whereas Woodbridge students are in the same school from grades K-8. Such school change has been shown to impact student performance. (Zhang and Zhang, Ref. 10). This does not seem to be a major factor in this case based on a comparison of student achievement results in the one other district in the state (Brandywine) in which relatively stable groups of students are in the same Grade 4/6 schools for both tests. Student achievement dropped in each of the three Grade 4/6 schools in Brandywine between Grade 4(2001) to Grade 6(2003). The average decrease in student performance and the increase in the achievement gap for the Brandywine district is comparable to that found for the state.

c). Because of its small size, the Woodbridge Science Specialist has been able to spend much more time mentoring and supporting K-5 teachers in their implementation of the science curriculum than have most of the other district specialists. This is certainly true for Brandywine where the Science Specialist serves both the three 4-6 schools and eight K-3 schools.

The superior results that Woodbridge obtained with their 2000 student cohort relative to the state averages were repeated in 2001. The matched student scores for the state from Grade 4 (2001) to Grade 6 (2003) showed a drop in average student achievement of 10.2%% and an increase in the achievement gap of 13.2%. In this same period, student achievement for Woodbridge students increased by 4.2% and the achievement gap dropped by 11.9%.

Further studies are in progress to better define the specific factors contributing to these differences in order to improve student achievement both at grade 6 and at grade 8.

TABLE 3

ACHIEVEMENT GAP CHANGE FROM GRADES 4 and 6 2000 to 2002				
% Grade 4 Students in PL 3,4,and 5 in Fall 2000 (5)				
	Woodbridge(1)	State(2)	State matched scores(3)	Brandywine
Student Demographics				
Number of Students	142	8274	7095	760
% Low Income	37%	36		34
% AA	44%	33%		39
Student Achievement				
% Total Students	82.4	85	87	83.6
% by race				
White	91.1	93.2	94.4	93.6
AA	67.9	70.6	73.7	68.4
Achievement Gap	23.2	22.6	20.7	25.2
% Grade 6 Students in PL 3,4,and 5 in Fall 2002(5)				
Number of Students	139	9036	7095	814
Student Achievement				
% Total Students	85.6	73.6	77.3	74.8
% by race				
White	90.7	85.2	86.8	89.1
AA	77.8	55.6	59.2	57.1
Achievement Gap	12.9	29.6	27.6	32
Change 2000-2002				
Achievement	3.2	-11.4	-9.7	-8.8
GAP Change	-10.3	7	6.9	6.8

Notes.

1. Woodbridge is a rural, low SES school district with one K-8 school.
2. Most students change schools between grade 4 and grade 6.
3. These compare valid scores of students who took both tests (Ref 9)
4. Brandywine is a large urban district with three Grade 4/6 schools. Most students are in the same school for both Grades 4 and 6.
5. Percent of students meeting/exceeding the Performance Standard

Significant increases in the Grade 8 achievement have been demonstrated through a professional development program for middle school teachers.

A pilot program with five school districts involving Grade 6-8 teachers from eight middle schools was held in early 2003 to test the impact of an intensive professional development using assessment of student work as a format. The initial sessions examined each school's student achievement data from the previous DSTP as a springboard for discussions on science assessment reform as outlined by the National Research Council (Refs. 11 and 12). Teacher questions developed from the data led to discussions of the depth of student understanding and how to measure it. All participating 6-8 teachers then administered the summative assessment developed for the unit they were teaching to their students. Scoring of these provided a framework for further discussions of the big ideas and central science concepts in the middle school curriculum, more authentic assessments for students and helped inform instructional improvement.

The amount of professional development and the process was structured to fit with each district's needs and time available. The dramatic results are summarized in Table 4.

TABLE 4
Raising Student Achievement in Grade 8 Science(1)

			% Students Meeting Standard		
PILOT DISTRICT	Schools	Professional Development	2002	2003	Change
Appoquinimick	2	18 hr+ follow up	45.7	68.1	22.4
Cape Henlopen	1	18 hr+ follow up	53.3	67.9	14.6
Capital	1	21 hours,	36.1	49.8	13.7
Milford	1	18 hrs	42.9	52.5	9.6
Colonial (Note 2)	3	9 hrs, during school day	48.3	55.7	7.4
PILOT (Average)			**45.2**	**58.8**	**13.5**
STATE(Ave)	33		43.9	49.0	5.0
REST OF STATE	25		47.1	49.3	2.2

Notes
1. Based on 2002 and 2003 Science DSTP results
2. Participation by teachers was voluntary.
Åll Gr 6-8 teachers participated in other districts.

Student achievement in pilot schools increased an average of 13.5% vs. 5.0% for the whole state and 2.2 % for the other 25 middle schools in the state. One district increased its scores by 22% while two others increased their scores by 14%.
These improvements in student achievement occurred with all races leading to a decrease in the achievement gap for the pilot schools of 6.2% compared to an increase of 1.1% for all other schools. Of special note, the achievement gap was reduced from 33.8% to 15.1% in the Appoquinimick district and from 46.6% to 28.9% in the Cape Henlopen district.

The improvements in student achievement and the reductions in the gap are clearly related to the intensity of the professional development program. The schools that had 18 hours of PD plus follow-up sessions and the school that had 21 hours of PD showed improvements of 22.4%, 14.6% and 13.7%. The schools in which teachers took fewer hours of PD and/or made teacher participation voluntary showed lower gains.

This is a very encouraging outcome. We believe that the discussions around the summative assessments help teachers focus on the big science ideas in the unit and the discussions on student work help teachers reflect on what student have and have not learned and inform their instructional strategies. This program is now being emulated with other districts while continuing to support further improvements in the pilot districts.

CONTRIBUTIONS: HUMAN RESOURCE DEVELOPMENT

One of the most important ways in which The Science Coalition has contributed to HR development is through the active program to develop teacher leaders who promote and foster implementation schools and districts.

Over 100 elementary lead teachers meet 4 days during the school year and one week in the summer to take professional development in science completing more than 7000 hours of professional development since 2001. In the early years of the LSCI, this professional development was primarily on content reinforcement. Since 1999, it has focused increasingly on all aspects of assessment by developing, piloting, scoring and disseminating summative assessments for every unit in the Grade 1-5 curriculum. These assessments are available to teachers through our website. The elementary Lead teachers are now beginning to work on formative assessments with a the goal of improving instruction by helping teachers understand and foster higher order thinking skills.

The Lead Teacher group has been the prime source for new Science Specialists, and a source of candidates for district administrative positions where their knowledge of science and their experience with the science program has helped strengthen support for science in the district.

At the middle school, 40+ middle school Lead Teachers meet two full days and three evenings each year to increase assessment literacy. During the first part of this project, they learned about assessment from experts, learned how to use assessments and then piloted a summative assessment with their students so they could learned to score student work, identify student misconceptions and plan strategies to deal with them.

The Science Specialists lead the program in their districts and have taken extensive professional development to strengthen their content knowledge and their leadership skills. As the project has expanded to middle and now is expanding to high school, these Specialists have been supplemented with Teacher-on-Loan Science Specialists working as part of the project leadership team.

The Core Evaluation concluded, *"one of the strengths of the LSC project consistently mention (sic) in all of the interviews and surveys was the establishment of the CSS and lead teachers to deliver professional development and follow-up support*"(Ref 3).

CONTRIBUTIONS: IMPROVEMENT OF EDUCATION INFRASTRUCTURE

The LSCI has led to major systemic changes in the education infrastructure in Delaware by creating and implementing a comprehensive program for systemic reform that engages all stakeholders in continuous improvement of science teaching and learning of science. The design and the success of the LSCI helped inform the proposal funded in 1999 by NSF for The Delaware Exemplary Mathematics Curriculum Initiative (DEMCI). The Social Studies community established a Social Studies Coalition in 2002 to foster similar collaboration and student improvement in Social Studies.

The specific infrastructure changes resulting from the LSCI include:

The Delaware Science Coalition

This Coalition fostered a collaborative culture in which district, state and higher education personnel and their business partners work together on the common goal of improving student achievement in science. It has also provided organizational continuity as the LSCI transitioned to the end of NSF funding.

With NSF funding, The Coalition was able to establish and continually improve a structured process that engaged districts in sharing rather than competing for resources, expertise and leadership. The key elements of the Coalition that have contributed to building this collaboration are:

- The Coalition Steering Committee in which representatives from all participating districts meet regularly to provide strategic program leadership.
- The Delaware Department of Education has supported and enabled many of the innovations introduced by the Coalition and provided leadership for the program.
- The Delaware Foundation for Science and Mathematics Education has been a valuable advocate with the business community and provided organizational expertise and critical funding.
- The Coalition Science Specialists whose major responsibilities are with their own district have worked together to provide the operational leadership required to manage the LSCI.
- The Lead teachers who lead science improvement in their school and have created a school based resource network to help each other.

A Comprehensive Model for Systemic Reform.

The LSCI program linked curriculum, materials support, professional development and assessment in a coherent structure with the goal of raising student achievement in science. As noted by our evaluators *" the LSC has operated in a seamless flow from standards, to professional development, to support for implementation, to student assessments which looped back into professional development."* (Ref 3)

A professional development system designed to prepare all K-8 teachers to meet the Delaware Science Standards.

The professional development process requiring 18 to 24 hours of professional development per curriculum unit was a major step change for Delaware teachers used to short "make–and–take" workshops. The professional development process gained teacher support by providing in-depth professional development from qualified instructors that was linked directly to the topics they must teach. Our evaluators noted, "*there is a highly shared vision of high-quality science education and professional development among the CSS who provide the professional development and the district administrators serving on the Steering Committee* "(Ref 3).

The professional development program has enabled schools throughout the state to prepare <u>and to maintain</u> the capacity of their elementary teachers to meet the science standards. The demand for elementary teacher professional development continued at the same level since NSF funding ended as it had in previous years; the demand for professional development of middle school teachers has doubled in the last two years.

A state supported Science Resource Center (SRC).

The SRC is the most visible and well-known symbol of the joint state, district and community commitment to improving student achievement in science. The SRC is *"a highly organized, well-run system maintaining, refurbishing and monitoring the use of 29 different kinds of kits for the whole "* that is *"one of "the strongest aspects of the LSC"* (Ref 3). The access to a wide range of curricular materials that the SRC provides to teachers and its dependability have been important in gaining teacher support for the science program both at the elementary and the middle schools.

The SRC has formalized a fee structure with the school districts and charter schools that will allow it to continue to provide both the materials and the professional development needed for all teachers of science. With state support, the SRC is being expanded to include laboratory facilities to better meet the needs for secondary teacher professional development in inquiry-science.

Continuing partnerships between the K-12 system, higher education and science based government agencies.

The LSCI established partnerships with the University of Delaware (U of D) and Delaware State University (DSU) from the very beginning of the program. UD and DSU faculty provided content courses to the Science Specialists in the early part of the LSCI when the entire focus was on elementary teachers. As the LSCI expanded to middle (and now to high) school, university faculty from both universities and experts from DNREC have partnered with CSS and other teachers to develop and improve curriculum and to provide professional development to K-12 teachers.

The Mathematics and Science Education Resource Center (MSERC) at the University of Delaware has been a critical partner in this effort. They have paired a "teacher in residence" as a Science Specialist with faculty both for the curriculum development process and for delivery of professional development. MSERC has worked with the School of Education to ensure that middle school teachers could use their courses for credit in a Masters of Instruction degree program. As part of an effort to build stronger curricular connections between science and mathematics, MSERC has developed and piloted a graduate course "Proportional Reasoning in Mathematics and Science" that will be available to all middle school Science Specialists and K-12 Mathematics Specialists. MSERC has also taken the lead in evaluating Lesson Study as an approach to improving teacher pedagogical content and thus enhancing student understanding with the support of an NCLB Higher Education grant.

A Sustained Science Program

The Science Coalition mission and activities have been successfully sustained surviving not only the end of NSF support, but also the retirement of both Principal Investigators. The Coalition program now is continuing seamlessly under the capable leadership of a new Science Associate.

All districts continue to support full time Science Specialists. More secondary specialists are being added to support the increased emphasis on middle and high school. Essentially all districts continue to participate in the Lead teacher program.

The Coalition now includes 5 Charter school and is actively expanding to high school under the guidance of the Coalition Steering Committee aided by a NSF pilot grant to begin the development of an inquiry based grade 9 and 10 curriculum and by continued support from the business community.

CONTRIBUTIONS: TO THE DISCIPLINE OF SCIENCE EDUCATION

The summative science assessments and the process for developing them have been a major learning experience that has contributed to better science education in Delaware and from which we believe many others working in the field of science education can learn. As outlined in the Outreach section, staff members from the Science Coalition have made a number of presentations at national and regional meetings on the topic of science assessments.

The process of preparing these assessments has resulted in:

- Improved teacher understanding of the central science concepts (big ideas) in each curriculum module.
- A much more consistent consensus among teachers across that state about what represents quality student work.
- Given teachers a diagnostic tool that provides fast feedback on student thinking and learning in their classrooms.
- Changed assessment from an appendage to the curriculum to the driver for what it means to "meet the standard".
- Helped shift professional development from an emphasis on the logistics and mechanics of teaching the lessons in the module to a much greater focus on instruction and student learning.

Rachel Wood (PI) and Dr. Julie Schmidt (Assessment Director) led the process with the active guidance and participation of Dr. Mary Ellen Harmon (Boston College) and consultations with Prof. Richard Shavelson (Stanford).

The key elements of the process that contributed to its impact are noted below. (The quotes are taken from a memo on The History of the Development of Delaware Comprehensive Assessment Program in Science (Ref 13).

- Development of clearly defined learning goals for each module required teachers to fully understand the central scientific concepts (the big ideas) in a unit and to articulate the relationship between the nationally developed curriculum materials and the Delaware Standards. If this showed, as it did in some cases, that the unit did not effectively express the big ideas, then the module and instruction needed to be modified to correct that gap. This not only increased teacher understanding, but also challenged their view that "teaching all the lessons in the module was tantamount to satisfying the standards".
- The process of clearly defining what each question measured "served as a validity check to ensure that we were measuring what we intended to measure by the question".
- The process of having teachers write their own answer to each questions led to unexpected disagreements that forced the grade level teams into intense discussions to reach consensus on what a complete response to a question should look like.
- Development of double digit scoring rubrics that includes both a score and a reason for the score "completely transformed our thinking and produced for us a paradigmatic shift in our understanding of assessment development. The rubrics could be written not only to evaluate achievement, but also to specify the various reasons that students were completely successful, all the different and common reasons they were partially correct, and distinguishes all the different reasons that students have something wrong in their thinking." We found that " making explicit an array of student thinking around a question forces teachers to think about the implications of instructional practice."
- Repeated field-testing of the assessments followed by revision of the questions, scoring criteria and rubrics based on student work. This iterative process was continued with each assessment until it satisfactorily measured what it was supposed to measure.

LESSONS LEARNED

1. Assessments will require as much reform as curriculum and instruction if science education is to be significantly improved. Assessments combined with student work are one of the most powerful tools for helping teachers understand the big ideas in the curriculum and engaging them in examining the impact of instruction on student learning.
2. Professional development of inquiry-based science usually begins with a focus on the kit materials and the logistics of classroom implementation. The focus needs to evolve to instruction and student learning in order to raise student achievement in the long run.
3. Higher education partnerships can benefit both K-12 teachers and college faculty. Teachers can get significant content reinforcement while the higher education partners can learn more the challenges of K-12 classrooms, effective instruction and student learning. The knowledge gained in such partnerships can facilitate much needed systemic changes in the preparation of pre-service teachers.

4. Partnerships with business are valuable as they provide access to important stakeholders who can be influential supporters and advocates for science both with the community and with education leadership.
5. Any Coalition involving many school districts must plan for high teacher and administrator turnover. Systems need to be established to ensure that teacher professional is readily available to all teachers and the partnership can survive the periodic loss of key district administrators.

PUBLICATIONS AND PRODUCTS

Websites

1. Delaware Science Coalition: www.sciencede.org
2. Science Assessments: www.scienceassessments.org
3. Delaware Science Page: www.k12.de.us/science

Presentations at National and Regional meetings

Staff of the Delaware Science Coalition have made a number of presentations at national and regional meetings and have assisted other science initiatives through discussions on the Delaware Science Coalition Program. Selected examples are listed below. (Note: Some of these meetings were partially supported by DuPont as part of their continued support of the Delaware Science Coalition.)

1. J. Coffey and K. Martin, Science Formative Assessments, NSTA, April 2004, (Atlanta).
2. K. Martin, G. Sharp, J. Wood, S. Wolford, A. McRae, Science Summative Assessments, NSTA workshop, April 2004.
3. K. Martin, G. Sharp, J. Wood, S, Wolford, A. McRae, Science Summative Assessments. NSTA Workshop, March 28,2003.
4. T. Meade and C. Fannin, Delaware Science Coalition Elementary Assessments, Washington State Teachers Meeting, Spokane, January 2004.
5. R. E. Wood, The Delaware Comprehensive Science Assessment. Workshop on Assessment in support of Teaching and Learning: Bridging the Gap Between Large-Scale and Classroom Assessment, National Research Council. Jan 23, 24, 2003.
6. R. Wood and K. Martin and a team of Science Specialists from the Delaware Science Coalition led a workshop on science summative assessments for Grades K-8 with particular attention to the use of double digit scoring rubrics. Tristate LASER Assessment Workshop, Newark, DE, Dec 9,10, 2002.
7. R.E. Wood and K. Martin, Symposium on Formative Assessments, NSTA April 2002.

Consultation with individual Science Initiatives.

1. R. E. Wood, Formative assessments, meeting with LSC OPA Hub, Clemson University, South Carolina, July 1999.
2. R.E. Wood, Science Assessment; Clemson University, July 2003.

3. Cincinnati Public Schools, Cincinnati, OH. C. Fannin, The Delaware Science Coalition Elementary Program, 2003
4. The Einstein Project*, Green Bay, WI. T Meade and C. Fannin, The Delaware Science Coalition Program, 2003,

Other Contributions

1. R.Wood served on the Committee on Classroom Assessment and the National Science Education Standards at the National Research Council that prepared the document Classroom Assessment and the National Science Education Standards., National Research Council Press, 2003.

2. R. Wood and J. Schmidt, History of The Development of the Delaware Science Coalition Assessment, personal communication to Dr. Lorrie Sheppard. Professor Sheppard featured examples from this document prominently in a chapter for Everyday Assessment in the Science Classroom, (Ref 13).

Local Outreach

J.W. Collette and H. Dillner, Progress in Improving Science Education in Delaware, Presentation to the Delaware American Chemical Society. September 2003.

References

1. Delaware Department of Education www.doe.state.de.us
2. Leadership and Assistance for Science Education Reform (LASER), National Science Resource Center, www.si.edu/nsrc/overv.htm
3. Research For Better Schools, Core Evaluation of Delaware LSCI (2002)
4. Research for Better Schools, Examining the Relationship Between a Science Professional Development Program and Student Achievement in Delaware Public Schools (2003).
5. L.Zhang, Examining the Consequences of a Standards-based Assessment Program from Student Perspective, National Council of Education Measurement, April 13-15, 2004.
6. Noble, Delaware Education Research and Development Center, University of Delaware, Personal communication (2003).
7. Diane J. Briars and Lauren B. Resnick, CSE Technical Report 528, Center for the Study of Evaluation, National Center for Research on Evaluation, Standards and Student Testing, UCLA (2000).
8. C. Dougherty, A Policymaker's Guide to the Value of Longitudinal Student Data, Education Commission of the States (2002).
9. L.Zhang, Delaware Department of Education, Personal communication (2004).
10. Y. Zhang and L.Zhang, Modeling School and District Effects in the Math Achievement of Delaware Students Measured by DSTP.
11. National Science Education Standards, National Academy of Sciences National Academy Press (1996).
12. Classroom Assessment and the National Science Education Standards, National Research Council, National Academy Press. (2001).

13. History of the Development of Delaware Comprehensive Assessment Program in Science, Rachel E. Wood and Julie Schmidt, Personal communication to Dr. Lorrie A. Shepard (2002).
14. Everyday Assessment in the Science Classroom, Janet E. Coffey, editor NSTA Press (2003).
15. R. Shavelson et al(2004), Personal communication.

APPENDIX A (Figures)

APPENDIX B

DELAWARE STUDENT TESTING PROGRAM (DSTP) -SCIENCE

The DSTP is a written assessment administered every fall to students in public schools in grade 4 and 6 and every spring to students in grade 8 and 11. The assessment was implemented for the first time in the fall of 2000 and is designed to measure student progress toward the Delaware content standards in science.

Success on the DSTP requires students to demonstrate their knowledge of, understanding of and capabilities in the content and skills articulated in the science content standards. The questions target student understanding of the big ideas in science including characteristics of living and non-living things, physical properties of materials, life cycles, structure/function, human body and healthy living, ecosystems, earth materials, earth in space, force and motion, and electricity.

The science test consists of two item formats: short answer (constructed response) and multiple choice. A sample of released science items as well as other information about the DSTP are available online at Delaware Department of Education website (**http://www.doe.state.de.us/aab/DSTP_intro.html**)

The data are reported as a Standards-based score (SBS) and Performance Levels (PL). The SBS measures students' progress toward meeting the Delaware Science Standard ad run from 100 to 500. PLs tell how the students are performing relative to the State's content standards. The five Performance Levels and the SBS are listed below.

SCIENCE

DSTP 1					
Grade	Well Below	Below	Meets	Exceeds	Distinguished
4	285 and lower	286-299	300-324	325-335	336 and higher
6	284 and lower	285-299	300-324	325-334	335 and higher
8	279 and lower	280-299	300-324	325-337	338 and higher
11	281 and lower	282-299	300-324	325-334	335 and higher

A comparison of the DSTP with TIMSS-R carried by R. Shavelson (Ref 15) has shown the following:

- The distribution of the types of knowledge (declarative, procedural, schematic) tested by the DSTP items is " strikingly similar" to that of TIMSS-R.
- There are systematic differences between the DSTP test and the TIMSS-R test in terms of the difficulty of the items in each of the classifications. The difficulty of

TIMSS-R questions increased from declarative to procedural to schematic; the DSTP followed the reverse order, but the differences were not statistically significant.

- The DSTP is 64% multiple choice, 36% open-ended questions whereas TIMSS-R is 72% multiple choice 28% open-ended questions