

# Standardized Test Outcomes of Urban Students Participating in Standards and Project Based Science Curricula

Robert Geier, Phyllis Blumenfeld, Ronald Marx, Joseph Krajcik, Barry Fishman & Elliot Soloway  
University of Michigan, 4027 School of Education, Ann Arbor, MI 48109  
Tel: (734) 904-2984, Fax: (734) 615-5245  
Email: bobgeier@umich.edu

**Abstract:** Considerable effort has been made over the past decade to address the needs of learners in large urban districts through scaleable reform initiatives. We examine the effects of a multifaceted scaling reform which focuses on supporting standards based science teaching in urban middle schools. The effort was one component of systemic reform efforts in Detroit Public Schools. Two cohorts of students that participated in reform curricula are compared with the remainder of the district population, using results from the high-stakes state standardized test in science. Both the initial and scaled up cohorts show substantive increases in science content understanding and process skills over their peers, and significantly higher “pass rates” on the statewide assessment up to a year and a half after exposure, with little fall off in achievement gains with increased scale. Examination of results by gender reveals that the curriculum reform effort succeeds in reducing the “gender gap” in achievement experienced by urban African-American boys. These findings are an encouraging example of how standards-based urban systemic reform can successfully impact student achievement as measured by statewide standardized instruments.

New standards based approaches to instruction present challenges to both teachers and students. For teachers using instructional methods based on recitation and direct instruction, inquiry teaching challenges them to develop new content knowledge, pedagogical techniques, approaches to assessment, and classroom management (Edelson, Gordin, & Pea, 1999; Hancock, Kaput, & Goldsmith, 1992; Marx, Blumenfeld, Krajcik, & Soloway, 1997). Inquiry learning challenges students too. It requires them to collaborate with peers, construct usable knowledge by linking new and old ideas, relate new science content to their lives in and outside of school, and self-regulate across the weeks that an inquiry project might unfold (Blumenfeld, Soloway, Marx, Krajcik, Guzdial & Palincsar, 1991; Krajcik, Blumenfeld, Marx, Bass, Fredricks, & Soloway, 1998).

Implementing standards-based instructional practice in large urban school systems presents a particular set of challenges for educators and their partners in reform efforts. These challenges include lack of resources, high levels of poverty, low student achievement, below grade level English proficiency, high student mobility, attendance problems, and difficulty recruiting and retaining highly qualified teachers (Lynch 2000, Hannaway & Kimball, 2001; Kahle, Meece, & Scantlebury, 2000). Despite these challenges, a number of researchers have shown that in highly resourced settings, inquiry instruction in urban classrooms can be successful (e.g. Bouillion & Gomez, 2001; Warren, Ballenger, Ogonowski, Rosebury, & Hudcourt-Barnes, 2001). The problem that remains, however, is to demonstrate that these efforts are scaleable beyond the design environments in which they were created to new sites with less concentrated resources and support, allowing the reform to progress successfully throughout an urban system.

In this paper we address this problem by presenting standardized test results gathered from a collaboration between the University of Michigan and The Detroit Public Schools as part of the Center for Learning Technologies in Urban Schools, funded by NSF. This collaboration was one part of a larger urban systemic reform effort. We demonstrate student achievement that can be attained when the focus of an urban reform is curricula aligned to standards and centered on specific learning goals. As Cohen & Ball (1999) recommend, the curricula are highly specified (the theoretical principles and methods are clearly defined) and developed (materials for teachers and learners are available and usable). Professional development is tailored to helping teachers enact the curriculum and changes in policy and management structures are sought to support the reform (Blumenfeld, Fishman, Krajcik, Marx & Soloway, 2000).

Previously, we reported that throughout each of four years of scaling, student pre-post test gains on curriculum specific assessments were significant, with generally large ( $>0.8$  SD) effect sizes (Marx, Blumenfeld, Krajcik, Fishman, Soloway, Geier, & Tal, 2003). Now we ask “Does the provision of standards based science instruction through project-based curricula lead to achievement gains on more distal standardized instruments?”

## **Background**

### **Teachers and Students**

During three years of implementation ending with the 2000-2001 school year, 37 teachers in 18 schools participated in our curriculum efforts, involving approximately 5,000 students. Across the district, 91% of the students are African American, and 69% are economically disadvantaged as measured by eligibility for free or reduced price lunch. Teachers and schools were selected based on several criteria, including adequate technology infrastructure, supportive administration, and equity among schools in access to innovative programs. Participating teachers averaged 11 years of teaching, of which 8 were primarily teaching science; approximately 70% report having a science credential. These figures were slightly lower than district statistics overall. (see Marx et al. 2003 for more details).

### **Curriculum Design**

The student population in this study had access to three curriculum projects developed by the University of Michigan in collaboration with Detroit Public Schools staff. Two projects were implemented in seventh grade (“What is the quality of air in my community?” and “What is the water like in my river?”) and one in the eighth grade (“Why do I have to wear a helmet when I ride my bike?”) The projects, which take about 8-10 weeks to complete, were designed to meet learning goals identified in Benchmarks for Science Literacy (AAAS, 1993) and the National Science Education Standards (NRC, 1996), and were aligned with the middle school science curriculum framework for Detroit Public Schools (Krajcik & Reiser, 2003; Singer, Marx, Krajcik & Clay-Chambers, 2000). Curricula were revised somewhat each year based on areas of weakness in student learning, along with teacher feedback and observer reports.

### **Learning Technologies**

Technologies are used in order to support the learning goals of the curriculum. Each software tool has been designed to take into consideration the unique characteristics of novice learners via embedded supports. Tools such as Model-It, Artemis, eChem, and data probes help expand the range of questions that students can investigate, the types of data and information that can be collected, and the types of data representations that can be displayed to aid interpretation. The tools are used across several curriculum projects and years, so that students become familiar with them and can benefit from repeated use. (Jackson, Stratford, Krajcik & Soloway, 2000; Krajcik, Blumenfeld, Marx & Soloway, 2000; Wu, Krajcik & Soloway, 2001).

### **Professional Development**

The goal of professional development is to prepare teachers to enact the curriculum appropriately for its underlying theoretical basis while adapting it to classroom circumstances. Professional development is conducted in summer institutes, monthly work sessions, teacher discussion groups and limited classroom support. The professional development effort makes use of student learning outcomes, content and pedagogical content concerns and teacher enactment difficulties among other topics (Fishman, Marx, Best, & Tal, 2003). In addition curriculum contain educative materials for teachers that focus on content, student understanding and potential enactment problems (see Schneider & Krajcik, 2002).

### **Michigan Educational Assessment Program Testing**

The Michigan Educational Assessment Program (MEAP) tests are a statewide standardized assessment aligned with the state objectives. The aggregate results for schools affect the statutory accreditation of the school and potentially the state funding allocation both directly and through enrollment pressures. Tests for middle school students are administered in January of 7th grade (math, reading) and in January of 8th grade (science, writing). The science test provides information on student understanding in three science content areas (Life, Physical, and Earth) and two areas of science process skills (“Constructing” and “Reflecting”). These five areas are combined to provide an overall science score, as well as a three-level student proficiency category which

divides students into two passing categories (“proficient” and “novice”) and one non-passing category (“not yet novice”).

## Methods

Through our partnership with the Detroit Public Schools, we obtained MEAP results for the school system from the 2000 and 2001 administrations of the test. The MEAP is subject to a variety of ongoing revisions, and the threshold scores for proficiency categories vary with restructuring and political pressures. Because of this variance and the yearly revisions of LeTUS materials, we chose to divide the analysis of MEAP findings by cohort. The division by cohort allows us to examine any change in effects from scaling and ongoing development efforts in LeTUS. Cohort I students may have participated in up to three LeTUS units: 7th grade Air Quality in the fall of 1998, 7th grade Water in the spring of 1999, and 8th grade Helmets in the fall of 1999. These students then participated in the administration of the MEAP test in January of 2000. Similarly, Cohort II students may have participated in the same three units the following years, culminating in MEAP administration in January of 2001.

The method of analysis involves a pooled comparison of students who participated in LeTUS curriculum with students in the Detroit Public Schools System who did not. It is important to note that this pooled comparison does not constitute a true control study since many of these students are experiencing other science improvement efforts as part of the NSF supported Urban Systemic Reform. We are aware of at least 6 other significant instructional interventions in the Detroit system. The comparison is thus between the LeTUS intervention and a combined pool of students receiving no intervention or other interventions in science. For the purposes of the initial analysis, we chose the most conservative criterion for distinguishing groups. Students were considered to have participated in LeTUS if they completed at least one LeTUS unit at any time during 7th or 8th grade year. We are thus examining whether participating in at least one project unit of 8-10 weeks has an impact on student understanding that persists for up to one year after the student completes the project unit. We also examine if the performance exceeds the impact on student understanding that results from regular instruction and other interventions in the remainder of the DPS population.

For the purposes of the analysis, "completion" of a unit was defined to include only those students for whom we had both start of unit (pre-test) and end-of-unit (posttest or student survey) data. Detroit Public Schools, like many urban districts, has a moderately high amount of student mobility and absenteeism, averaging approximately 20% in our curriculum enactments. This definition of completion therefore excludes students who may have participated in part or all of the curriculum, placing them in the district-wide comparison group. This gives us a treatment sample of 760 students in Cohort I and 1,043 in Cohort II, with a comparison groups of 8,900 and 8,662 respectively.

In addition, we conduct similar analyses disaggregating the effect of LeTUS participation by gender. Achievement results for largely African-American urban males are of particular interest because of the at-risk nature of this population in Detroit and other urban districts. Because of the highly charged nature of racial achievement reporting, we did not have access to student level data on ethnicity. Given that the student population of Detroit exceeds 91% African American and is less than 3% white, we do not feel there would be sufficiently representative samples to allow for a meaningful comparison in any event.

## Results

Our first analysis compares the students who completed at least one LeTUS unit with the general pool of DPS students who did not participate in a LeTUS curriculum. The total raw MEAP science score is examined, along with the raw score totals for each of the 5 content and process subscores, using a standard comparison of means (Student's t). Table 1 shows the results and the relevant effect sizes for the Cohort I and II students. All differences were significant at  $p < .001$ .

In Cohort I, students who completed at least one LeTUS unit during 7th or 8th grade significantly outperformed their DPS peers in all content and process categories measured by the science MEAP. The effect sizes are respectable, with participation in at least one unit corresponding to a 14% improvement in total score when compared with the remaining DPS population. The raw score difference of approximately 50 points corresponds to answering 5 additional multiple choice questions correctly (out of 54) or receiving full credit for

2.5 additional free response questions (out of 6). This amounts to a standardized effect size of 0.44 standard deviations. Participation in a LeTUS unit contributed most strongly to Earth Science content scores; earth science is a well-integrated component of the 7th grade curriculum units in particular.

Cohort II represents a substantial upward scaling of the intervention to include more school sites and instructors, with consequently less individualized support for teachers in the classroom. As with the Cohort I group, the second cohort demonstrates significantly higher achievement in all five content and process categories measured by the MEAP science assessment compared with the other students in the district. Participation in at least one LeTUS unit corresponds to 40 raw score points, the equivalent of 4 additional multiple choice questions or 2 free response questions answered for full credit. This amounts to a 13% difference and a standardized effect size of 0.37 standard deviations. Overall and individual category effect sizes are slightly but not appreciably attenuated with the scale-up, with the exception of the physical science content category. Interestingly, the reduced effect size for physical science corresponded to a significant curriculum revision in our 8<sup>th</sup> grade physical science unit, with which some teachers reported difficulty.

**Table 1. MEAP Science Comparison: LeTUS Cohorts vs. DPS Population**

	<b>Cohort I</b>			<b>Cohort II</b>		
	LeTUS Mean	DPS Mean	Eff. Size (std.)	LeTUS Mean	DPS Mean	Eff. Size (std.)
Sample Size	<i>n</i> =760	<i>n</i> =8900		<i>n</i> =1043	<i>n</i> =8662	
Total Score	389.16	340.40	0.44***	360.05	320.03	0.37***
<i>Content Areas</i>						
Life Science	75.09	67.73	0.28***	97.40	85.99	0.32***
Physical Science	83.80	73.67	0.36***	52.28	49.31	0.17***
Earth Science	66.01	53.15	0.53***	89.55	77.67	0.34***
<i>Process Areas</i>						
Constructing	77.25	67.34	0.34***	53.49	47.73	0.29***
Reflecting	87.00	78.52	0.33***	67.33	59.33	0.32***

\*\*\* indicates statistical significance (independent samples Student's *t* test at  $p < .001$ ).

## Pass Rates

Because of the threshold nature of determining passing categories, the gains from exposure to LeTUS curriculum are more dramatic when one examines MEAP passing rates for students. The MEAP passing categories are the "high stakes" numbers for students and school districts. Figure 1 below compares MEAP passing rates to the general DPS passing rate for both cohort years. Participating in at least one LeTUS unit is associated with a 19% increase in passing rate in Cohort I and a 14% increase for Cohort II. The differences are statistically reliable (Chi Square 117.8, 103.1;  $p < .001$ ). These gains in MEAP passing rate have very important positive implications for the schools and the district, beyond the learning gains for the individual students.

## Gender Differences

The Detroit Public School district-wide MEAP results show statistically significant gender differences between girls and boys achievement on the science assessment. Boys score 17 points below girls in 2000 and 14 points below in 2001 (effect size of 0.15 SD,  $p < .001$ ). These follow the trend that has been widely reported in urban schools with predominantly African-American populations, where the boys under perform in comparison to the girls (see Kahle et. al. 2000).

Our data, however, suggests that at least for the population of students in LeTUS program schools, participation in at least one LeTUS unit has an attenuating effect on this gender difference in achievement. Figure 2 shows the interaction of gender with participation in LeTUS curriculum by cohort. In both cohorts there is an apparent reduction in boy-girl achievement differences on MEAP associated with LeTUS participation. The interaction term is marginal for the first cohort, but reaches conventional levels of statistical reliability for Cohort II (2 way ANOVA, interaction  $F=4.59$ ,  $p < .05$ ).

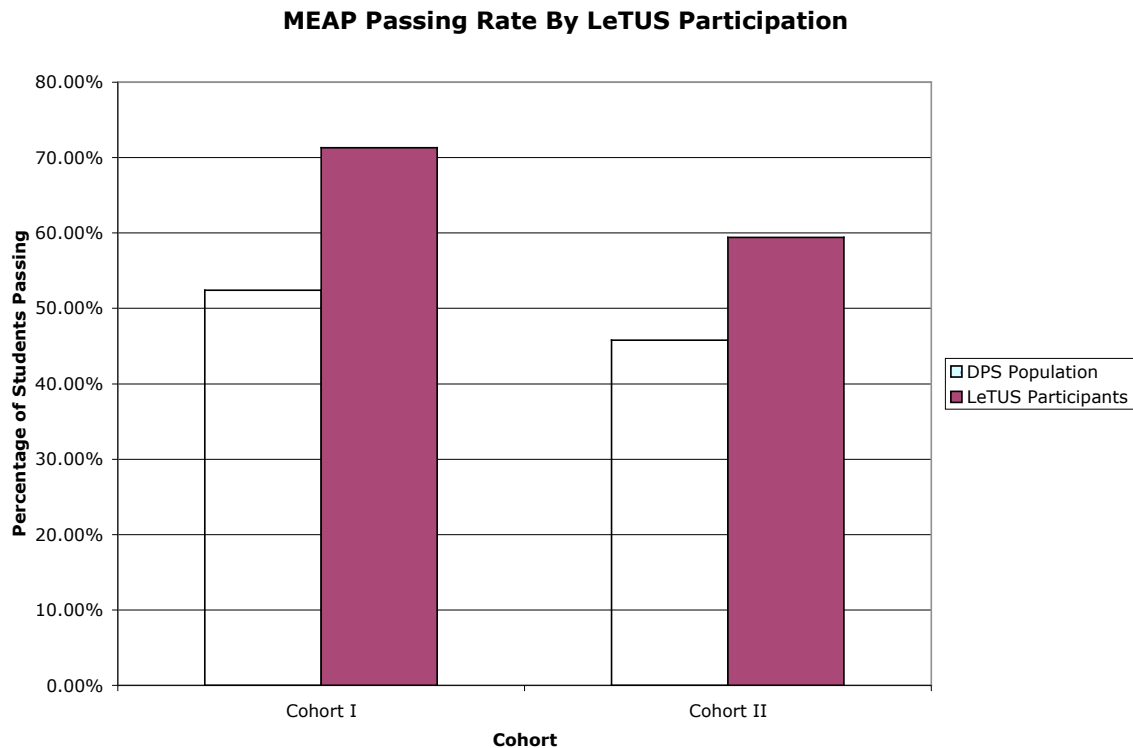


Figure 1. MEAP Passing Rates by LeTUS Participation

As with the overall score results, the trend toward reducing the performance gap between boys and girls is more dramatic when we examine the MEAP proficiency categories considered “passing.” Table 2 shows MEAP passing categories for each cohort. In each case, there is a substantial difference between students who had completed at least one LeTUS unit and those who did not. While both boys and girls benefit from participation, the gender gap between passing rates is largely attenuated for LeTUS participants in Cohort I. In Cohort II we see that the boys involved in at least one LeTUS unit “caught up” and showed no significant difference from the girls.

This is an intriguing and very positive finding. Standards-based instruction including pervasive technology and project-based curricula appear to engage at-risk urban male learners, narrowing and closing the gender gap in achievement with their female peers.

Table 2. MEAP Passing Rate by Gender (LeTUS Program Schools)

	Girls		Boys	
	No LeTUS	LeTUS	No LeTUS	LeTUS
Cohort I Passing Rate	51%	75%	47%	71%
Cohort II Passing Rate	52%	59%	43%	60%

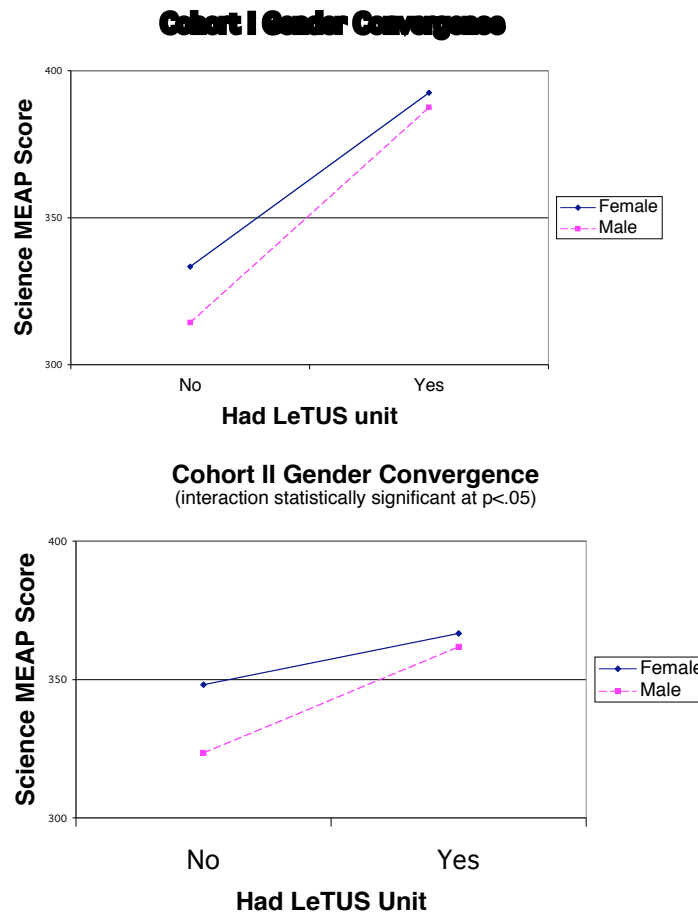


Figure 2. Gender Interaction with LeTUS Participation on MEAP Score, Cohorts I and II.

## Alternative Explanations

We examined several potential confounding effects and alternative explanations for our findings, including missing data, school, student, and teacher selection bias. Given student mobility and absenteeism, one possible explanation for our results is that weaker students dropped out during participation in a LeTUS unit or showed a higher absenteeism rate and were excluded from the LeTUS sample by not being present for a pre- or post-test. This might result in classifying more academically able students as LeTUS participants yielding a net positive effect. While it was unlikely that this group was sufficiently large to affect the results in an analysis involving ten thousand students, we repeated the analysis for each cohort, this time including all students who had any contact with a LeTUS unit for any length of time, so as to eliminate any attrition effect. The results presented were unchanged.

In our first effort to rule out school selection bias, we removed from the non-LeTUS sample all of the students from DPS schools that were not involved in LeTUS project enactments during the period in question for each cohort. The analysis was then repeated for each cohort, comparing LeTUS participants with non-LeTUS participants in the same schools. As with the analysis of student attrition, the results of this comparison were largely unchanged, though the effect sizes for Cohort I were slightly stronger than in the original analysis (stats). We further checked for school selection bias by examining the mean MEAP scores for the 8th grade non-science tests in writing, comparing buildings participating in LeTUS unit(s) with the remainder of the DPS population. Again, we found no significant differences nor evidence of school selection bias.

To rule out student selection bias we made use of the 7th grade MEAP administration in mathematics and reading, as a proxy for prior academic achievement. As with the 8th grade tests, the 7th grade MEAP is administered in late January. Unfortunately, this falls after participation in the fall LeTUS unit on air quality, so it is not a true prior ability measure. Mathematical topics like graphing and data interpretation and reading skills like identifying arguments are covered in the LeTUS Air Quality unit. Our analysis found a slight positive bias in students' 7<sup>th</sup> grade MEAP scores favoring students who completed at least one LeTUS unit. The largest effect size for this effect was .03 standard deviations. Given the very small effect size and the additional possibility that the scores were influenced by the already completed Air Quality unit, we do not find evidence to suggest that our general results were a product of bias in student selection.

A final alternate hypothesis is that some bias in teacher selection contributed to our findings. Teachers who chose to participate in reform efforts might tend to be mavericks or show greater levels of commitment to self-improvement, bringing a broader repertoire of techniques to the classroom. Since the nature of our intervention was targeted at teacher practice and measures of prior teacher performance are not available, we have no analytic method to rule out this possibility. However, the nature of the teacher selection process did not lend itself to such a bias since participants were chosen on the basis of schools. Additionally, pre-test to post-test gains on our "close" curriculum specific tests showed high teacher variability in outcomes, suggesting that if there were a selection bias it would likely be relatively small. Our teacher pool was comparable to the general Detroit teacher population in highest degree earned, and slightly weaker than the DPS average in terms of teaching outside of certification area and years of experience.

## Discussion

The partnership between the Detroit Public Schools and the Center for Learning Technologies in Urban Schools offered a unique opportunity to construct a multifaceted reform effort that focused on highly specified, well-developed educative curriculum supported by ongoing professional development, embedded technology, and policy-level assistance in implementation. The collaboration also allowed us to investigate achievement outcomes on distal state standardized tests, using student-level data to isolate the effects on individual students who participated in LeTUS curriculum enactments. This level of alignment and data availability can be difficult to obtain in reform initiatives that encompass multiple districts or states, where partnerships may be more difficult to develop, test instruments and data formats may vary, and local control of curriculum and policy may be zealously guarded.

In the context of the educational challenges of a large urban district, our results are in many ways quite heartening. Students who were exposed to one or more LeTUS units during 7th or 8th grade significantly outperformed their DPS peers in all content and process categories measured by the science MEAP. The comparison group itself was engaged in a number of reform efforts as part of the Detroit urban systemic efforts; our results reflect improvements in achievement beyond those from the district-wide interventions. The effect sizes for the differences averaged approximately 0.4 standard deviations, and were reflected in 14-19% increases in MEAP passing rates.

Our efforts support and add to the work of other groups who have examined student-level distal achievement data in conjunction with standards-based reform efforts. While some of the larger statewide or multi-state studies have found no effect from standards-based teaching after controlling for other factors (eg. Berends, Chun, Schuyler, Stocly, Briggs, 2002), other researchers have been able to identify gains associated with the reform efforts and standards-based instruction, but often with relatively small effect sizes between 0.05 and 0.15 (Klein, Hamilton, McCaffrey, Stcher, Robyn, and Burroughs, 2000; Kahle, Meece, and Damjanovic, 1999; Borman 2002). When the nature of the reform allows for the implementation of specified and well-developed curricula to support other aspects of the reform effort, larger effect sizes have been achieved (Ross, Sanders, Wright, Stringfield, Wang, & Alberg, 2001; Borman 2002). Our results confirm that a multi-faceted approach to reform incorporating highly specified and developed curricula can lead to substantive increases in student achievement.

Moreover, the substantial improvement of boys' achievement and pass rates is somewhat surprising. It is congruent with reports by Kahle et al (2000) that attitudes of urban minority males are more positive when their teachers engaged in standards based professional development. Our speculation is that in contrast to

traditional instruction, the variety of experiences provided by these curriculum, such as collaboration, inquiry, technology use, are more appealing to boys and therefore result in greater participation and learning.

## References

- American Association for the Advancement of Science. (1993). *Benchmarks for science literacy, Project 2061*. New York: Oxford University Press.
- Berends, M., Chun, J., Schuyler, G., Stockly, S., Briggs, R.J. (2002). Challenges of Conflicting School Reforms: Effects of New American Schools in a High-Poverty District. Santa Monica: RAND.
- Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling-up technology-embedded project-based science in urban schools. *Educational Psychologist*, 35(3), 149-164.
- Blumenfeld, P. C., Soloway, E., Marx, R. W., Krajcik, J. S., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, 26(3&4), 369-398.
- Borman, G. D., Hewes, G. M., Overman, L. T., Brown, S. (2002). *Comprehensive School Reform and Student Achievement: A Meta-Analysis* (CRESPAR-R-59). Maryland: Publications Department, CRESPAR/Johns Hopkins University, 3003 N. Charles Street, Suite 200, Baltimore, MD 21218.
- Bouillion, L.M., & Gomez, L.M. (2001). Connecting school and community with science learning: Real-world problems and school-community partnerships as contextual scaffolds. *Journal of Research in Science Teaching*, 38, 878-898.
- Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement* (CPRE Research Report Series RR-43). Philadelphia, PA: Consortium for Policy Research in Education.
- Edelson, D. C., Gordin, D. N., and Pea, R.D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences* 8(3&4) 391-450.
- Fishman, B., Marx, R., Best, S., & Tal, R. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, 19(6), 643-658.
- Hancock, C., Kaput, J.J., & Goldsmith, L.T. (1992). Authentic inquiry with data: Critical barriers to classroom implementation. *Educational Psychologist*, 27, 317-364.
- Hannaway, J. & Kimball, K. (2001). Big isn't always bad: School district size, poverty, and standards-based reform. In S. H. Fuhrman (Ed.). *From the capital to the classroom: Standards-based reform in the states*. 100<sup>th</sup> Yearbook of the National Society for the Study of Education (Part II). Chicago: University of Chicago Press.
- Jackson, S. L., Stratford, S. J., Krajcik, J., & Soloway, E. (1994). Making dynamic modeling accessible to precollege science students. *Interactive Learning Environments*, 4(3), 233-257.
- Kahle, J. B., Meece, J. L., & Damjanovic, A. (1999). *A pocket panorama of Ohio's systemic reform, 1999*, [Brochure]. Oxford, OH: Miami University.
- Kahle, J. B., Meece, J., & Scantlebury, K. (2000). Urban African-American middle school science students: Does standards-based teaching make a difference? *Journal of Research in Science Teaching*, 37(9), 1019-1041.
- Krajcik, J., Blumenfeld, B., Marx, R. and Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Minstrell & E. Van Zee (Eds.) *Inquiry into inquiry: Science learning and teaching*. Washington: American Association for the Advancement of Science Press.
- Krajcik, J., Blumenfeld, P., Marx R. W., Bass, K. M., Fredericks, J., & Soloway, E. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Science*, 7(3&4), 313-350.
- Krajcik, J. & Reiser, B (2003). Design Principles for Developing Inquiry Materials with Embedded Technologies. In Marx, R *Partnerships for Urban Systemic Reform: The Effects of Inquiry Curriculum Developed by the Center for Learning Technologies in Urban Schools*, Symposium conducted at the Annual Meeting of the American Educational Research Association, April 21 - 25, Chicago, IL.
- Lynch, S. J. (2000). *Equity and Science Education Reform*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers, 10 Industrial Avenue, Mahwah, NJ 07430.
- Marx, R. W., Blumenfeld, P., Krajcik, J. S., & Soloway, E. (1997). Enacting project-based science. *Elementary School Journal*, 97(4), 341-358.
- Marx, R., Freeman, J., Krajcik, J., & Blumenfeld, P. (1997). Professional development of science teachers. In B. Fraser & K. Tobin (Eds.), *International handbook of science education*. Dordrecht, The Netherlands: Kluwer.
- Marx, R., Blumenfeld, P., Krajcik, J., Fishman, B., Soloway, E., Geier, R., & Tal, T. (2003). Inquiry based science in the middle grades: Assessment of learning in urban systemic reform. *Journal of Research in Science Education*, under review.
- National Research Council (1996). *National science education standards*. Washington, DC: National Academy Press
- Ross, S.M., Sanders, W.L. Wright, S.P., Stringfield, S., Wang, L.W., and Alberg, M. (2001). Two- and three-year achievement results from the Memphis Restructuring Initiative. *School Effectiveness and School Improvement*, 12(2):323-346.
- Schneider, R.M. & Krajcik, J. (2002). Supporting science teacher learning: The role of educative curriculum materials. *Journal of Science Teacher Education*, 13 (3), 221-245.
- Singer, J., Marx, R. W., Krajcik, J., & Clay-Chambers, J. (2000). Constructing extended inquiry projects: Curriculum materials for science education reform. *Educational Psychologist*, 35(3), 165-178.
- Warren, B., Ballenger, C., Ogonowski, M., Rosebury, A., & Hudicourt-Barnes, J. (2001). Rethinking diversity in teaching science: The logic of everyday sense-making. *Journal of Research in Science Teaching*, 38, 529-552
- Wu, H.-K., Krajcik, J. S., & Soloway, E. (2001). Promoting conceptual understanding of chemical representations: students' use of a visualization tool in the classroom. *Journal of Research in Science Teaching*, 38, 821 - 842.

## Acknowledgments

Funding for these projects was made possible by the Spencer Foundation in a grant for "Technologies to Enable Inquiry: The Influences on Student Learning and Motivation" and by the National Science Foundation under the following programs: REPP (REC-9720383 REC-9725927 REC-9876150) and USI (ESR-9453665).