Participants: Yacov Avichai, Bret Feranchuk, Lourdes Monteagudo, Armando Triana (TAMS), Henry Frisch (U. Chicago), Dave Hedin (Northern Ill.)
+ part-time HEP faculty, postdocs, staff, and graduate students.
+ 1 Masters student (Northern Ill.)
+ 4 undergraduates (2 U. Chicago),
+ 2 High School interns (1 U. Chicago)

0.1 Introduction

This MRC has two major components, with the first focussed on quantitative statistical assessment of interventions in math and science education (Subsection 0.2). This component has three distinct yet closely related strands: 1) assessment of program impact on students, 2) determining readiness of schools for intervention, and 3) quantitative determination of program factors in successful schools. The second component of this MRC involves developing a model for the transfer of these and other research results to the daily working and policy-making of individual schools, an area that has received little attention but one we believe is critical (Subsection 0.3). Lastly, synergies with the other MRC's in this proposal are discussed in Subsection 0.4.

0.1.1 Background

Ten years ago, The Teachers Academy for Mathematics and Science (referred to as TAMS or the "Academy") was conceived and initiated by physicists and local parents and educators. Initially funded by the NSF and DOE, the Academy is now fully supported by the State of Illinois and private funds at a level of more than six million dollars annually. The Academy seeks to train a critical mass of elementary school teachers in mathematics and science, as well as parents and the leadership team (e.g. principals, local school board members, etc.) at a large number of individual schools. The Academy does this by implementing a comprehensive, school-wide, three-year program that consists of three major components for teachers.

First, the program involves classroom instruction addressing both math and science content as well as teaching pedagogy. This is done using methods appropriate for teachers who instruct elementary level students (In Chicago these teachers are non-specialists; 17,000 teachers regularly teach math and science in the Chicago Public Schools, very few with any training in math and science). Second, extensive classroom support in each individual teacher's classroom is provided, including long-term follow-up with weekly visits to the classroom by TAMS specialists.. This assists teachers in transferring their increased knowledge into practice in their own classrooms. And finally, the Academy distributes appropriate classroom teaching materials and student manipulatives, including access to computer technology. A more detailed description of the Academy's program can be found on the TAMS web site at http://www.tams.org.

By the State's request, TAMS has now extended its work from Chicago to schools in other areas of Illinois including, among others, Joliet and Aurora, which are in close proximity to Argonne and Fermi National Laboratories respectively. To date, the Academy has worked with 114 elementary schools and 3,000 teachers in Chicago and four other urban centers in Illinois. The Academy purposely chooses to work with schools who teach an ethnically diverse population of students (currently 61.0% of the students in Academy schools are African-American, 30.3% Hispanic, 6.1% Caucasian, 2.6% Asian, and 88.8% receive reduced priced lunch, a common measure of low-income families.).

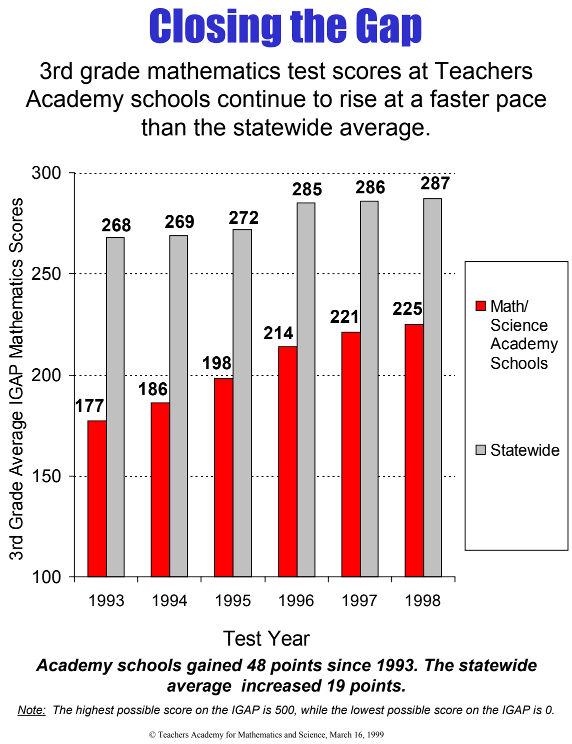

Figure 1: A comparison of 3rd grade mathematics test scores at TAMS schools versus the statewide average.

0.1.2 Program Results

Student Results Urban schools in Illinois, as well in as the nation as a whole, face diverse and complex challenges[1][2][3]. Improvements in the situation require that the problems can be broken down and approached systematically. The more widespread use of standard-based tests and better data sets will offer the opportunity to attack these challenges through rigorous statistical methods.

One measure of student achievement in Illinois is the Illinois Goal Assessment Program (IGAP). The IGAP is a standardized test based on goals set by the Illinois State Board of Education. Elementary students are tested in third, sixth, and eighth grades for mathematics, and fourth and seventh grades for science. Scores are reported by school and grade as average scores[1] on a scale from 0 to 500. A sampling of the Academy's results are shown below. Teachers Academy schools, while starting at a lower initial point, have improved substantially more rapidly than other schools statewide. On average from 1993-1998, 3rd grade mathematics students at Academy schools have gained 48 points (from xxx to xxx) compared to 19 points on average (from xxx to xxx) for other schools in the state.

It is important to note that, by design, the Academy works with some of the lowest performing schools in the state where educational support mechanisms are most needed.

In contrast to the data in Figure 1 graph, for which the total number of schools which had participated in the mathematics and science programs was held constant over the years, the graph in Figure 2 includes schools in three-year comparisons tabulated by when they entered the Academy.

[1] The numbers given in the following charts are average scores/school.

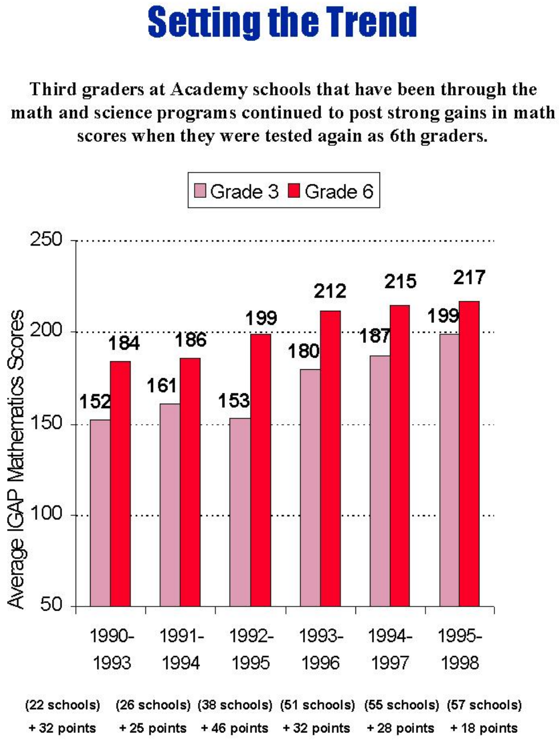

Figure 2: The average IGAP scores of students who have gone through the math and science programs in TAMS schools tested in the 3rd grade and again in 6th grade.

All comparisons: 1990-1993, 1991-1994, 1992-1995, 1993-1996, 1994-1997, and 1995-1998 indicate sustained improvement over time and show marked, positive differences in comparison to state third to sixth grade increases

over the same time spans (the corresponding state gains are 4, 2, 4, 3, 4, and 3 points respectively). These comparisons are statistically significant, 1994-1997 ($p<.01$) and 1995-1998 ($p<.05$).

Seventh graders at Academy schools out-gained their peers at other schools statewide by 29 points on average from 1993-1998. Last year's (1997-1998) gain was most pronounced with 19 points versus the statewide gain of eight (8) points. Both comparisons are statistically significant at the $p<.01$ level.

Much in the same fashion as the mathematics results, science scores show substantial gains from fourth grade to seventh grade. The largest gain experienced was in the most recent time period from 1995 to 1998. In 1995, when Academy school fourth graders were tested they scored an average of 159 points. Three years later, in seventh grade, they gained 48 points to score an average of 207 points. The same comparison for the state yielded a ten (10) point gain, from xxx to xxx.. This difference is statistically significant ($p<.01$).

In addition to the overall average, individual Academy schools showed impressive double-digit gains from fourth grade in 1995 to seventh grade in 1998. Out of 45 schools with 4th and 7th grade scores available, 43 experienced double-digit gains. The other two schools had double-digit losses for the same time period. In comparison, 70% of the other statewide schools showed double-digit gains, 10% single-digit gains, 19% losses, and 1% no change. Overall, six of the top 25 gaining schools in the state out of the 645 in this three-year comparison were Academy schools.

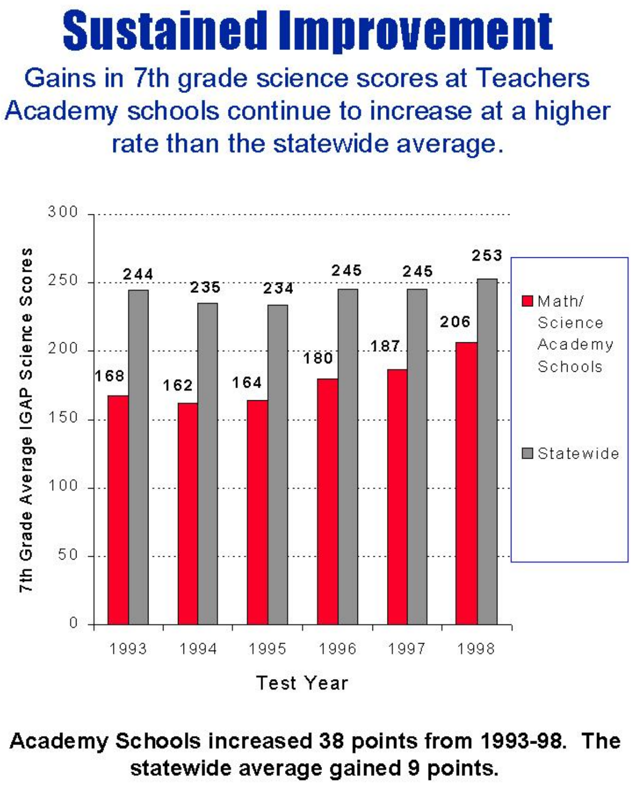

Figure 3: IGAP science scores for 7th graders at TAMS schools and statewide as a function of year tested.

Teacher Results In addition to gains in student achievement, teachers who have been exposed to the Academy's professional development program have improved their own knowledge. A 29-item open-ended basic math and science skills test, with a maximum total of 40 points, was administered before and after the intervention. Inferential comparisons were made only for the teachers who had a matched set of data (i.e. both a pretest and post-test). Because this comparison was made only for a subset of teachers, it is important to compare mean scores for all participating teachers for the pretest (mean = 11.4), and the post-test (mean = 22.9) to determine if the comparison data were biased due to sampling attrition. The means for these larger groups were slightly smaller than the mean scores for those teachers where a pretest–post-test comparison was possible (although within measurement error parameters). Teachers in the East St. Louis School District showed a 98.3% improvement on the skills tests, while teachers in the Joliet School District showed a 32.3% gain in one year. Briefly, the following tables highlight some of the gains for the two districts as well as the distribution of scores[2]:

[2]More detailed statistical results from this study, or any of the analyses mentioned in this proposal, can be obtained from the Academy. Note pretest and post-test totals might not add to exactly 100% due to rounding.

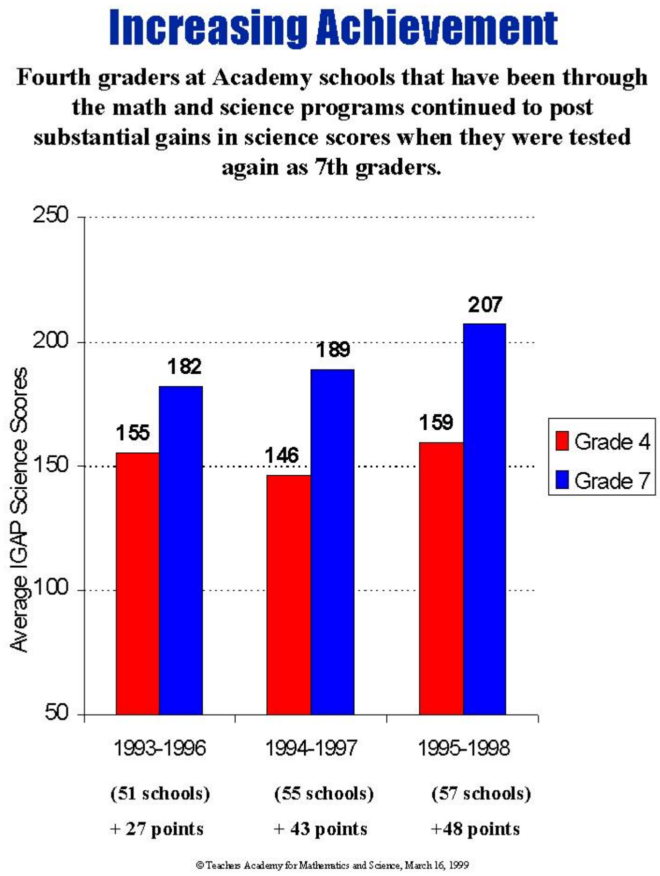

Figure 4: Average IGAP scores for three cohorts of 4th graders who have been through the TAMS math and science programs, tested in the 4th grade and again as 7th graders.

0.1.3 Overview of the Proposed Research: The DataSets

Inspired by these preliminary results, TAMS has recently gained access to large data sets made available by the Boards of Education of Chicago and the State of Illinois in order to investigate more rigorous methods of educational research. These data sets encompass approximately 6 million student records, as well as data on more than 3,000 schools, over 10 years. These data sets contain student test score data, school and student demographic information, and teacher statistics. This information will be combined with internal Academy databases involving programmatic data such as teacher attendance rates in the Academy's program, teacher and student content knowledge and attitudinal measures, and both quantitative and qualitative classroom observations. This component's approach echoes the call of The President's Council of Advisors on Science and Technology (PCAST) which stated, " (a) large-scale, rigorously controlled program of research and evaluation will ultimately be necessary"[4]. More detail is provided below in Section 0.2

Teachers and school administrators rarely have the expertise, tools, or personnel to help them analyze data[7]. As a result they rely on intuitive decision-making much more often than other professionals do in similar sized organizations[8]. This leads to the second major component of this MRC, namely the development of models for transfer of research results into tangible efforts at individual schools. The focus of this component is supported by recent reports of The National Academy of Evaluation (NAE)[5] and the National Research Council (NRC)[6] whose strategic goals call for: 1) tying

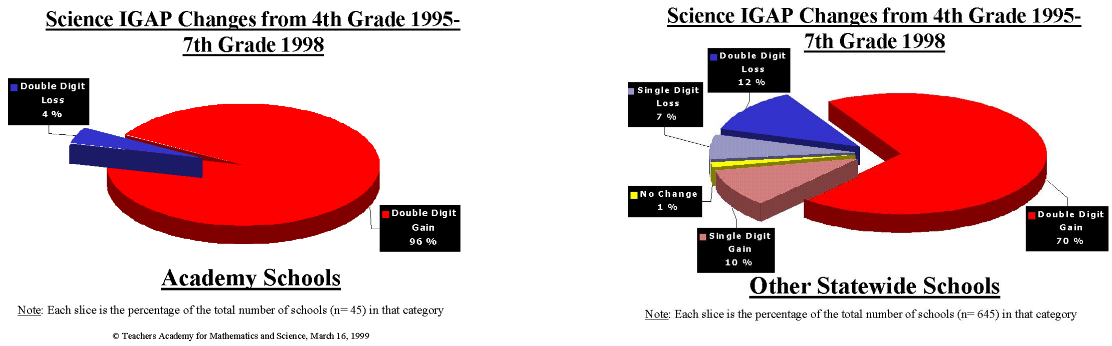

Figure 5: The distribution in changes in the average science IGAP scores between the 4th grade in 1995 and the same students in the 7th grade in 1998 in TAMS schools (left) and other schools statewide (right).

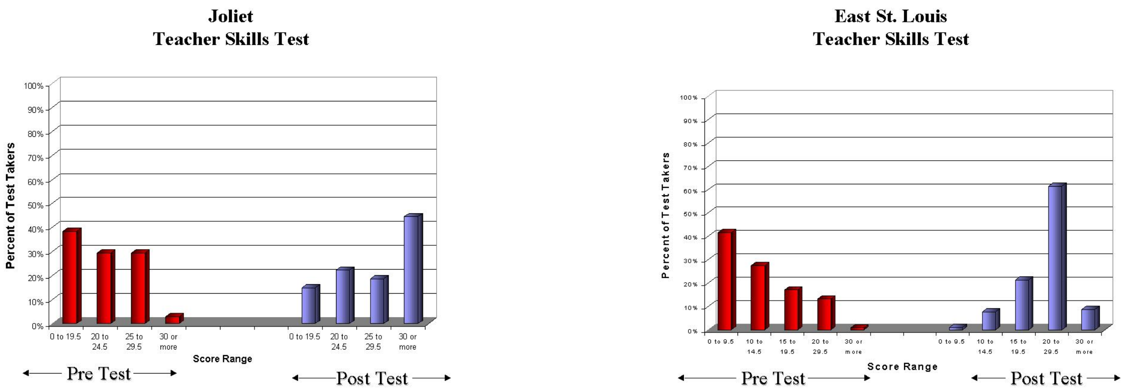

Figure 6: Pretest and post-test scores on a teacher skills test for teachers from East St. Louis (left) and Joliet (right).

"research to practice"; 2) focusing on "critical transitions in the lives of low-income students"; and 3) "linking teachers' professional development and teaching practices with improved student learning". This component will also be described in detail below, in Section 0.3

The funding of this MRC will help pay for: 1) the preparation of the large data sets for statistical analysis, 2) the analysis of these data along multiple research strands, 3) the dissemination of the results at the city, state, and national levels, and 4) the transfer and use of the knowledge gained in a manner useful to educational practitioners in Academy schools.

0.2 Assessment of the Effectiveness of Interventions in Math and Science Education

The large datasets, once they have been prepared for analysis, will allow the quantitative investigation of a number of questions. We discuss the three we have selected in turn below.

District/Score Ranges	Pretest	post-test
East St. Louis[3]	Number of Teachers	Number of Teachers
0 to 9.5	41.5%	1.0%
10 to 14.5	27.4%	7.7%
15 to 19.5	17.0%	21.2%
20 to 29.5	13.2%	61.5%
30 or more	0.9%	8.7%
Average Score (40 max.)	11.9	23.6
Joliet[4]		
0 to 19.5	38.2%	14.8%
20 to 24.5	29.4%	22.2%
25 to 29.5	29.4%	18.5%
30 or more	2.9%	44.4%
Average Score (40 max.)	21.4	28.3

Table 1: Scores from a Pretest/Post-test comparison for teachers from East St. Louis and Joliet who had completed training with TAMS.

0.2.1 Impact of Intervention on Student Performance

One research strand of this MRC is to develop rigorous and sophisticated statistical models for measuring the impact of interventions within schools. One of the biggest challenges of this type of research involves addressing the question of how the impact of an intervention can be measured in the absence of adequate comparison groups, which is often the case (for example, in the case of TAMS, the worst schools in the city were chosen for intervention, making the selection of a control sample problematic). With the assistance of the former Corporate Statistician for the CNA Insurance Companies, Yakov Avichai, the Academy has begun evaluating the impact of the Academy's work with teachers using techniques from the insurance industry to solve the problem of not having adequate control samples. This is particularly important for the Academy since its target population consists of the poorest performing schools and those with high concentrations of low income students[9]. The goal has been to analyze the impact of the Academy intervention on 7th grade science score differentials on state standardized student test scores between 1994 and 1999.

Ideally, one would select a set of comparable schools to those that were exposed to the Academy's program and proceed with a straightforward comparison. This, however, is not feasible given the state of available information (and sample sizes) and is doubtfully achievable in the foreseeable future. Instead, we have developed a multivariate regression model of all schools that controlled for key independent variables. The multivariate fit enabled us to calculate the expected results for Academy schools in the general environment and compare them with the actual results. Thus, we were able to document a 15% gain versus comparable schools attributable to the Academy's intervention while controlling for different socioeconomic conditions. Additionally, positive effects were seen in 91% of the participating Academy schools. We should point out that this methodology can be modified and adapted to work for assessing any intervention, being in no way tied exclusively to the specific TAMS programs. In addition, it allows determination of which factors are important to gains in student achievement, even in the face of high mobility, broad socioeconomic variation (e.g. Hispanic vs Afro-American populations), and other factors that make the usual assessment so difficult.

As part of the proposed research we will explore the nature and causes of the random variability that is the background for student performance data. We will treat the identified background effects explicitly in our models. In this manner this strand can be thought of as analogous to several of the other MRC's in this proposal which propose to isolate "interesting" signatures from among much background data. This proposal will allow for the extension of these models to include additional independent variables such as program attendance rates, number of teachers receiving the program, and duration of the program as well as to individual student-level as opposed to the current measure of school-level performance. Additionally the proposed research will examine the effects on students of having certain patterns of well-trained elementary teachers. For instance, does having a few well-trained teachers in earlier years have a sustained effect in future years? Is this effect greater than having well-trained teachers later in a student's education? Does having several well-trained teachers in consecutive years produce a cumulative effect on student performance? The answers to these questions can have important policy implications in decisions involving where to apply resources in order to achieve the greatest impact.

0.2.2 Determination of School Readiness for Intervention

The allocation of limited resources is an important problem in math and science educational reform interventions. A second research strand of this MRC is development of mathematical models which address the question of how to select schools in order to best utilize the limited resources available. This question is important because there are both direct and indirect costs associated with selecting for an intervention or program schools that for a variety of reasons may be unreceptive or uncooperative. In addition to the direct financial costs to the program, there are also indirect human costs associated with selecting unreceptive or uncooperative schools. These costs involve the loss of program services and benefits to equally needy children in more receptive environments. When limited resources are spent in unreceptive settings, these children are denied access to the potential benefits the program can provide their schools. In effect, this line of research addresses the readiness question, that is, which schools are not currently ready, at this point in time, to benefit from the intervention program. We note that this is a problem that does not occur for the more common small pilot intervention programs, but is a problem that needs to be addressed in any program committed to systemic change in a large school system.

The TAMS program is an intensive, whole-school activity. The literature on successful professional development and organization change is clear in this area[10]schools need to understand the purpose of the professional development activities and trust its representatives so that they will "buy-in" to the process. There are several stages of preparedness and readiness for school change and its essential that the school's stage matches that necessary for the provider's program[11]. The first part of the TAMS intervention program is to work with schools to bring them to a developmental stage where the school could benefit from the Academy's program. It is important to note that while the Academy is committed to working with schools that are among the most challenging (i.e. schools with high concentrations of low income students, historically underprivileged groups, and low levels of student achievement), unfortunately this group of potential clients is quite large (there are 600 public elementary schools in Chicago alone). Thus, it is impractical to reach all of them at any one time, and this necessitates choosing some over others. The proposed research addresses the very practical question of how to do make school selection decisions in a systematic fashion to minimize risk to Academy resources while still selecting schools with extensive needs. Early efforts along this research strand appear promising.

0.2.3 Successful School Factors

The third research strand of this MRC involves quantitative identification of program factors that make schools successful. This involves quantifying the requisite level of teacher participation. There are two elements to this quantification. The first will quantify teacher participation in terms of number of teachers at an individual school to receive the program (i.e. the critical mass). The second will quantify teacher participation in terms of the number of hours of the program attended by each teacher (i.e. intensity) over a given period of time (i.e. duration).

This strand of the proposed research will allow us to directly address a number of important programmatic issues and questions. For example, how much teacher participation is needed to affect student test score performance? That is to say, does reaching 50% of total teachers in a school make a difference in student performance? Or does one need to reach 75% of teachers in a school? Or does one need to involve even a higher percentage? How much of a difference in student performance do these different levels of participation produce? Additionally, the proposed research will allow us to examine how successful one-on-one visits to teacher classrooms are in increasing students' academic performance as measured on standardized tests? The answers to these questions are important because they are generalizable to other educational professional development initiatives, and because they have serious implications for scalability and reproducibility of math and science programs in other geographical locations. Hence, the results of this strand of the MRC's proposed research should be beneficial to policy makers and educational practitioners alike.

0.3 Building Schools' Internal Capacity for Data-driven Improvement

In order to induce change in educational systems it is crucial to translate beneficial research into effective practice. A second component of this MRC will be focusing on timely and expeditious transfer of research results to individual school communities. Teachers and school administrators rarely have the expertise, tools, or personnel to help them analyze data[7]. As a result they rely on intuitive decision-making much more often than other professionals do in similar sized organizations[8]. As part of the proposed MRC the Academy will initiate a pilot project in 3 schools. The objective of this pilot project will be to implement a process that will assist schools in the development of their own internal capabilities to use data for decision-making.

Many educators still have little experience using and looking at data. Principals and teacher leaders will be assisted by the Academy's School Data Specialists in how to use data to assess needs. The pilot project will also be designed to facilitate an inquiry cycle using the newly acquired data sets; as well as assist the principals, teachers, school board members, and other school leaders in using data to develop school improvement efforts that are tied directly to teaching and student achievement. This will give principals and teachers the opportunity to development their ability to gather, analyze, and use data in an appropriate manner. The process[12] will begin with the development of a 'dynamic school profile'. This profile will contain templates, tools and processes that can be used by schools to guide their improvement efforts. It will contain guiding questions and specific elements, trends and needs that staffs and stakeholders could look for in the data. As a result, schools will develop an effective school improvement plan, and an increased capacity to continually use data to improve student achievement. From the results we achieve in the pilot, we will seek funds to modify and adapt this into a well-documented and publicly available program scalable for a larger number of schools, supported by State, other Federal (non-JCET), or private funds.

0.4 Synergies

The statistical tools, presentation and interpretation of data, and the infrastructure necessary for handling of large data sets have much in common with those used in physics data analysis. In fact, the availability of powerful computing facilities available to us through several of the other JCET institutions will give us the ability to create the initial master data sets much faster and more efficiently than on our own. In addition, University of Chicago and Fermilab physicists have been closely involved with this effort through the TAMS Program Committee, on which they serve. This close cooperation has already proved to be highly profitable for both those that come from an education background and from a physics background. Some of the regular seminars of JCET will be devoted to statistical and educational topics; the exposure of young physicists (including the undergraduates at the JCET institutions) to this will, in our experience, induce further involvement and outreach in their subsequent careers. Additionally the Academy works with schools in locations geographically close to all the universities and national laboratories involved in this collaboration allowing for interaction between the physics and educational components of the Center. The Center will also be an arena in which visitors from all over the country will be introduced to effective math and science education and cutting-edge educational research.

0.5 Summary

In summary, the research strands of this MRC address questions of how to measure the impact of a math and science program in a rigorous statistical manner, how to select potential sites while minimizing the associated risk, and how to quantitatively determine the essential factors in a school necessary for systemic improvement in student performance in mathematics and science. Additionally, this MRC addresses the need to transfer this research to practice and build research capacity within individual schools themselves. The Academy believes that, taken as a whole, this research will make a substantial and unique contribution to the educational research literature as well as to practitioners, policy makers, and the public at large. Furthermore, it provides a unique opportunity to involve research scientists and science students, both graduate and undergraduate, in the realities of the complex challenges (and possible solutions) of math and science education in urban elementary schools.

References

[1] Cole, C., Flores, C., Medley, C., Merchant, B ., Miron, M., Sloane, F., Stake, R., *The Burden: Teacher Professional Development in Chicago School Reform*. Chicago: CIRCE, University of Illinois at Urbana-Champaign (1994).

[2] Bryk, A. *et al.*, *A View from the Elementary Schools: Reform in Chicago*, Chicago: Consortium on Chicago School Research, University of Chicago (1993).

[3] Hess, F., *Spinning Wheels: The Politics of Urban School Reform*, Brookings Institution: Washington DC, (1999).

[4] "Report to the President on the Use of Technology to Strengthen K-12 Education in the United States", President's Council of Advisors on Science and Technology, Washington, DC (1997).

[5] "Investing in Learning: A Policy Statement with Recommendations on Research in Education by the National Educational Research Policy and Priorities Board", National Educational Research Policy and Priorities Board, U.S. Department of Education, Washington, DC (March 1999).

[6] "Improving Student Learning: A Strategic Plan for Educational Research and Its Utilization", National Research Council, National Academy Press (April 1999).

[7] Killion, J. and Bellamy, T., *Journal of Staff Development*, **21**:1 (Winter 2000).

[8] Berman, P. and McLaughlin, M., *Federal Programs Supporting Educational Change*, Santa Monica: The RAND Corp., (1977).

[9] Feranchak, B., Avichai, Y., Langworthy A., Triana A., "Evaluation of a Mathematics and Science Professional Development Intervention: A regression approach to comparing a subgroup of schools that is neither representative nor random", Paper scheduled to be presented at the Annual Meeting of the American Educational Research Association, Seattle, WA (2001).

[10] Fuller, M., *Change Forces*, Bristol, PA: Falmer Press, Taylor-Francis (1993).

[11] Schmoker, M., *Results: The Key to Continuous School Improvement,* 2^{nd} *edition*, Association for Supervision and Curriculum Development: Alexandria, VA (1999).

[12] Langworthy, A. and Feranchak, B., "Data-Driven School Improvement: Creating a Dynamic School Profile, Paper scheduled to be presented at the Annual Meeting of the American Educational Research Association, Seattle, WA (2001).