INTO THE FRAY: NOVICE TEACHERS TACKLE STANDARDS-BASED MATHEMATICS

B. SCHRODER and S. GNINGUE
Dept. of Middle & High School Education, Lehman College, CUNY
Bronx, NY 10468
barbara.schroder@lehman.cuny.edu , serigne.gningue@lehman.cuny.edu

Abstract

This article tracks twenty-one graduates of a reform-based mathematics teacher education program for two years as they begin teaching mathematics in public elementary schools in New York City. Using surveys, classroom observations, and interviews, it examines the extent to which these beginning teachers were able to implement standards-based mathematics instruction in their classes. Results of the study were mixed. The novice teachers generally demonstrated an adequate understanding of the underlying mathematics principles and strong intentions of teaching mathematics for understanding. They were generally able to engage children in learning, and most performed at the "beginning stages of effective instruction" in their first year. However, they still struggled to engage students in higher order thinking and knowledge construction. In their second year their abilities improved, but they were still hampered by local factors such as insufficient in-service support, the restrictions of high-stakes testing, and the overall school climate.

Introduction

The National Council of Teachers of Mathematics (NCTM) envisioned a classroom where students learn mathematics with conceptual understanding [1]. In this vision, the teacher engages students in critical, in-depth, higher order thinking about the content through the use of manipulatives, technology, cooperative learning, and other pedagogy that enable students to construct mathematics concepts on their own by reasoning, verifying, comparing, synthesizing, interpreting, using different strategies of investigating or solving problems, making connections, communicating ideas, and constructing arguments.

However, despite the NCTM's recommendations, there is still a wide variation in content coverage and teaching practice among teachers. Many teachers are comfortable with teaching only how to perform calculations without meaning or context. Liping Ma's study of elementary teachers' understanding of mathematics in China and the United States found that participants in the United States viewed mathematics as "an arbitrary collection of facts and rules in which doing mathematics means following set procedures step-by-step to arrive at answers" [2]. Further, prospective elementary teachers often expect to implement traditional teaching practices, such as teacher-directed lessons followed by student practice of paper-and-pencil computations, aimed at developing procedural fluency. For Timmerman, these practices send students a clear message that mathematics often does not make sense, and they do not need to or cannot understand mathematics [3].

In elementary schools, the initial focus in teaching mathematics concepts is on whole numbers built upon "counting" processes, and students do not face many cognitive difficulties until they start moving from the whole number set to the rational set. The transition to the rational set that involves complex conceptual changes unfortunately functions as a "gatekeeper," keeping many children from advancing in their understanding. Earlier representations that worked for whole numbers are inadequate for understanding concepts in the rational number system such as fractions. Rational numbers can be variously understood; e.g., as fractions, or ratios of whole numbers, as measures of length or area, as operations involving stretching/shrinking, or as extending the integers via multiplicative inverses. These numbers may be variously represented, such as equivalence classes of fractions, "mixed numbers," finite or repeating decimals. Each of these mathematical ideas is complicated, and a large body of research in education and cognitive science (such as, "Rational Number, Ratio, and Proportion" by M.J. Behr, et al.) addresses them [4]. Fractions provide a basis for understanding and manipulating rational expressions that occur later in algebra, units of measurement fundamental to science, and probability/statistics. The transition from the whole number system to the rational system is therefore critical, making learning for understanding crucial at the elementary school level, especially in urban schools, where many children struggle to achieve in mathematics.

Yet, the teaching of mathematics in elementary school is still geared to memorizing routine skills with little room for reflection and logical thought. For many teachers, doing mathematics is associated with following the teacher's rules. "Knowing mathematics means remembering and applying the correct rule and having the answer ratified by the teacher" [5]. This leads students to an "unfortunately limited picture of mathematical expertise," in which to be an expert, one merely needs to be able to remember and explain the correct rule and have the answer ratified by the teacher [6].

This situation is exacerbated in urban settings where most minority children live and attend school. Ball, Hill, and Bass found that in the third grade for instance, "...higher-knowledge teachers tended to teach non-minority students, leaving minority students with less knowledgeable teachers who are unable to contribute as much to students' knowledge over the course of a year" [7]. As a result, substantial numbers of urban students do not attain the mathematics skills and understanding needed for success in today's world. Other factors, which range from high teacher turnover and dilapidated schools to bureaucratic inertia, crime, drugs, and poor housing, contribute to urban students' failure in mathematics. One of the main problems, however, remains the manner in which mathematics concepts are being conveyed to students.

The consequences of this situation at the national level are that the United States faces not only a shortfall of mathematicians, but a general mathematics illiteracy. The number of mathematics Ph.D.s granted today is two-thirds the number awarded in the 1960s, and student achievement is below that of other countries. By the time students reach their college mathematics courses, they are bored, passive, hostile, and full of complaints [8].

Context of the Study

In efforts to address these issues, teacher education programs at the City University of New York (CUNY) have developed courses over the past ten years that use inquiry-based approaches to instruction, in which students have opportunities to construct their own understanding of basic concepts. These efforts were supported during 1996-2000 by the National Science Foundation's (NSF) Collaboratives for Excellence in Teacher Preparation (CETP) program, which brought together faculty members from both education and mathematics/science departments across the country to redesign core offerings for prospective elementary teachers. In most instances, pre-service teachers taking the redesigned courses were engaged in activities they could apply directly in their own classrooms. The faculty involved in these courses anticipated that they would result in graduates who would be prepared and willing to implement these same sorts of practices in their own classrooms.

The data examined here were collected as part of an evaluation study of this NSF-funded program, the New York Collaborative for Excellence in Teacher Preparation (NYCETP). The larger evaluation study focused on the teaching of both mathematics and science, and sought to evaluate the impact of NYCETP through, among other methods, collecting evidence of the classroom use of standards-based teaching strategies by teacher graduates of the CETP-reformed pre-service classes through classroom observations, surveys, and interviews.

Purpose of the Study

This article examines the classroom practices of twenty-one novice elementary teachers in New York City public schools, graduates of one CUNY college, as seen through the lens of the evaluations designed for the CETP study. The overarching purpose is to ascertain the extent to which they were able to incorporate the skills of and understandings about teaching mathematics acquired in their pre-service classes into their first two years of teaching.

Research Questions

The study's central question is: "To what extent were these teachers teaching for conceptual understanding?" Other questions are listed below.

- Were there significant changes between pre-service teachers' intentions in using inquiry-based teaching practices and their first-year self-reports of practice?
- To what extent did CUNY elementary education students provide their students with the opportunity to engage in inquiry-based mathematics activities in their first year of teaching?
- As teachers moved into their second year in the classroom, were there significant changes in their overall performance, with regard to teaching mathematics to K-6 schools?
- What factors seemed to support or detract from their teaching for conceptual understanding?

Subjects

Data were collected over a four-year period. In 2001-2002, prospective participants in the evaluation study were surveyed in the last year of their pre-service teacher education program concerning, among other things, their intention to implement a series of standards-based reform practices in the classroom, all of which were aimed at promoting students' conceptual understanding.

In 2002-2003, among the larger pool of participants were twenty-one first-year teachers and student teachers who were observed teaching a mathematics lesson. In addition to being observed in the classroom, participants completed surveys and submitted sample lesson plans. Throughout this article, the term "Year I" refers to 2002-2003.

In 2003-2004, eleven of these teachers continued in the study, and were observed a second time teaching a mathematics class. As in the previous year, they completed surveys and submitted lesson plans, and their principals completed surveys. Throughout the article, the term "Year II" refers to this year.

Finally, in 2004-2005, six teachers participated in intensive interviews, in which they reflected upon various aspects of their experiences in the mathematics classroom the previous year.

Instrumentation

The surveys, classroom observation protocols, and lesson plan scoring rubrics were originally developed by the University of Minnesota for the CETP core evaluation study, and were either used as is or modified slightly. The interview protocol was developed by the CUNY Center for Advanced Study in Education. The instruments used are outlined below.

Teacher Surveys — Teacher surveys were administered once before participants began teaching, and then again each time that they were observed teaching. The initial survey, adapted from the CETP K-12 Teacher Survey, was administered in participants' pre-service education classes. It asked them to predict to what extent they expected to implement a variety of standards-based teaching strategies, for example asking students to write descriptions of their reasoning or to work on problems related to real world or practical issues, in their own classrooms. Then, either right before or immediately after they were observed teaching, they completed another survey that among other things, asked them to relate the extent to which they were actually implementing these same strategies in the classroom.

Classroom Observations — Classroom observations were the essential element in determining teachers' levels of teaching for conceptual understanding. Participants were observed by faculty members of their college as they taught a mathematics lesson, using a Classroom Observation Protocol (COP) created for the national CETP evaluation study. The COP "constructed through the selection of items from several classroom observation forms. The items selected were those that had been shown to be predictive of standards-based instruction and positive student outcomes. Sources used included: Horizon Research, Inc.; the Arizona Collaborative for Excellence in Teacher Preparation; Evaluation of the Long-Term Effect of Teacher Enhancement project; the Constructing Physics Understanding evaluation project and the Systemic Initiatives evaluation project" [9]. Preliminary estimates of the instruments' inter-rater and intra-rater reliabilities range from 50% to 80%, while internal consistency analyses resulted in alphas of .90 or better.

The first three sections of the COP described the teacher's background, classroom demographics, and the classroom context. The fourth section collected information in the areas of, among others: **Type of Instructional Activities** (e.g., lecture, small group discussion, hands-on activities, etc.); **Levels of Student Engagement** (e.g., high, medium, or low); and **Level of Student Cognitive Activity** (e.g., "receipt of knowledge," "construction of new knowledge"). A fifth section rated **Key Indicators** of standards-based instruction, such as "The lesson encouraged students to seek and value alternative modes of investigation or problem solving," and asked the observer to comment on the effectiveness of the lesson in several categories. Finally, the sixth and last section contained an overall **Capsule Rating** of the lesson's effectiveness, ranging from 1 (ineffective instruction) to 8 (exemplary instruction).

The CUNY faculty observers were all seasoned mathematics educators, school administrators and/or teacher educators. In order to promote inter-rater reliability, two training sessions were held: a full day attended by observers across all the campuses, and a half day on each individual college campus. In these sessions, observers rated sample videotaped lessons and then compared their ratings. When differences were observed among raters during these training sessions, they were only on the order of one out of a possible seven in the vast majority of cases. Observers then discussed their differences in order to resolve them and arrive at a shared understanding of the ratings.

<u>Lesson Plans</u> — The teachers submitted a sample lesson plan, usually for the observed lesson. The plans were rated by teams of two faculty members, who discussed and resolved their rating differences for each plan. The rating protocol focused on essentially the same standards-based elements as the observation protocol. Ranging from 1 (not evident) to 4 (clearly evident), the ratings indicated the extent to which the lesson plans "encouraged students to manifest characteristics of students in standards-based classrooms." More specifically, they were asked if they provided students with the opportunities to: 1) engage in in-depth, higher order thinking about the content; 2) make real-world connections; 3) work together as learning community; 4) feel as though they have their needs met; and, 5) learn important concepts. A final question (6) asked the extent to which effective questioning strategies for constructing student knowledge and differentiating instruction was present. As with the Classroom Observation Protocol, raters also provided a **Capsule Rating** of the planned lesson's effectiveness ranging from 1 (ineffective lesson) to 8 (exemplary lesson).

<u>Interviews</u> — Six of the teachers participated in in-depth interviews concerning their experiences teaching mathematics during the winter of 2004. The intent of these interviews was to provide

additional understanding of how teachers evaluated their students' math achievement, as well as of the school contexts within which they operated. The interview protocol asked teachers to describe their classes, in general terms and in terms of mathematics instruction. It also asked them to rate their students' mathematics performance and describe their reasons for these ratings.

The pages that follow first summarize the data from the teachers' first year of teaching. We then move on to consider the experiences of the eleven teachers who were observed again teaching a mathematics lesson in their second year. The article ends with a discussion of the role of the contexts within which they taught, drawing on the surveys and interviews.

Results — Changes between Pre-Service Intentions and First-Year Self-Reports of Practice

According to the NCTM in 1991, "The constraints of the real world of schools overwhelm the perceptions these new teachers hold about what mathematics teaching and learning could be [and] ...the result is that many new teachers find it difficult to adapt what they have learned in their teacher preparation programs" [10]. The results of this study indicate that the situation has not changed significantly as of 2002. When surveyed in their university education classes prior to their observations, prospective teachers indicated a strong commitment to employing the classroom strategies consonant with helping children develop a good conceptual understanding of mathematics. When asked how often they intended to implement a list of twenty-two reform practices (e.g., "have students use or make conceptual or mathematical models") with possible responses ranging from 0 (never) to 3 (regularly), the majority of respondents chose either "2" or "3" for most items (mean of 2.45 across all items).

In their first year of teaching, however, the reported use of reform strategies dropped dramatically from the stated intentions of the previous year. The average responses dropped for all items, leading the overall average to drop significantly from 2.45 to 1.46, which when analyzed with a paired-samples, T-test analysis, showed a statistical significance of .000. Listed below are some illustrations of such decreases.

- Eighty-one percent of respondents intended to "have students write descriptions of their reasoning *regularly*" before beginning teaching, but only 20% reported actually having them do this *regularly* in their first year of teaching.
- Sixty-three percent of the respondents intended to "have students make connections to other fields and areas of mathematics *regularly*," but none reported doing this *regularly*, and only 33% reported doing it *occasionally* in their first year of teaching.
- Sixty-three percent intended to "engage students in data collection and analysis *regularly*," but only 7% reported doing it *regularly* in their first year in the classroom.

Results—Year I Lesson Plans

For five of the six indicators listed above, the percent receiving a rating of 1 or 2 was at least two times the percent rated at 3 or 4. The mean capsule rating was 2.86, slightly below a rating of 3 or "elements of effective instruction." Taken together, these results indicate that overall, the raters did not see in these lesson plans strong evidence of the ability of the participants to plan for standards-based mathematics instruction.

Results—Classroom Observations

The new teachers were observed teaching a mathematics lesson using the Classroom Observation Protocol (COP) developed by the University of Minnesota. For each class, the observation time was divided into five-minute intervals, during which the observer recorded the Type of Instructional Activity, the Level of Student Engagement, and the Level of Cognitive Activity seen. The lessons observed ranged in length from twenty minutes to one hour, with the average length being forty minutes, or eight five-minute segments. In total, 164 five-minute segments were coded and analyzed.

Classroom Instructional Activities — Nineteen different types of instructional activities that could take place in a mathematics classroom were identified in the COP. For this study, we grouped them into four categories: Type I, those deemed typically traditional (e.g., those that were definitely not standards-based, such as "lecture" or "reading seat work") were observed in forty-two of the 164 five-minute segments, or 26% of the time; Type II, activities in between "traditional" and "standards-based" (e.g., problem modeling or demonstrations) that could be used during both standards-based and traditional instruction were observed 31% of the time; Type III, inquiry- or standards-based activities, such as collaborative learning groups, were observed 37% of the time; and, Type IV, occasions when teaching was interrupted due to student behavior, administrative tasks, or other non-related learning tasks were observed 6% of the time.

Student Engagement — Three levels of student engagement were identified in the COP instrument: high (when at least 80% of students actively participated in the lesson); mixed; and, low (when at least 80% of students were off task). Of the 160 five-minute intervals with student engagement data, students were highly engaged in 58, or 36% of the time. They were somewhat engaged 49% of the time, and only minimally engaged 15% of the time.

Cognitive Activity — Five levels of cognitive activities were coded in the observations. Level 1 (Receipt of Knowledge) was noted when students were involved in the reception of information;

Level 2 (Application of Procedural Knowledge) when students applied their knowledge by doing worksheets, practicing problems, or building skills; Level 3 (Knowledge Representation) when students manipulated information by reorganizing, categorizing or attempting to represent what they learned in a different way; Level 4 (Knowledge Construction) when students created new meaning by making connections, generating ideas or solving new problems; and, Level 5 (Other) for administrative tasks, interruptions, etc. Of the 164 five-minute recordings available, Level 1 and Level 2 were recorded 84% of the time, meaning that students spent most of their time receiving and applying procedural knowledge.

Relationships Among Lesson Elements — To examine whether there were patterns of relationships among the three lesson elements recorded by observers, we first converted Type of Instructional Activity into a scale variable: a value of 1 was assigned to Type I activities, a value of 2 was given to Type II activities, and a value of 3 was given to Type III activities. Similarly, Level of Student Engagement was converted into a scale variable ranging from 1-3, and Level of Cognitive Activity was converted into a scale variable with values from 1 – 4. (It was possible for the observer to enter more than one Type of Instructional Activity into one five-minute segment. When this happened, such values were first averaged across all entries in each five-minute segment of the observation and then averaged across the entire set of 164 five-minute segments observed.)

Table 1 presents the means of the resulting measures. As seen there, observers rated the lesson taught highest in the area of Student Engagement (on average 2.4 out of 3) and lowest in the level of student Cognitive Activity (on average 1.69 out of 4). Several Pearson's Product-Moment Coefficients were then calculated, to explore the bivariate relationships among the three variables. Table 2 shows that all three correlations can be considered statistically significant at the .05 level. The strongest relationship observed was that between Type of Instructional Activity and students' Cognitive Activity, and the weakest was between Type of Instructional Activity and Student Engagement.

Table 1
Coded Classroom Instructional Activity, Student Engagement and Student Cognitive Activity Ratings, Year I

Variable	Range	Mean
Type of Instructional Activity (reform value)	1 - 3	1.99
Level of Student Engagement	1 - 3	2.40
Level of Cognitive Activity	1 - 4	1.69

Table 2
Type of Instructional Activity, Student Engagement, and Student Cognitive Activity: Correlations

	Pearson's Product-Moment Coefficient	*p*
Type of Instructional Activity with Student Engagement	.213	.036
Type of Instructional Activity with Cognitive Activity	.581	.000
Student Engagement with Cognitive Activity	.312	.002

Ratings of Key Indicators — In the fifth section of the COP, the observer was asked to rate each of nine "key indicators" using a Likert scale ranging from 1 to 5 ("not at all" to "to a great extent"). The overall mean across the nine indicators was 2.55. This result seems to suggest that the standards-based characteristics were observed to a minimal extent. The four lowest ratings indicate:

- Students were not reflective about their learning. Seventy-one percent of the observations were given a 1 or 2 for this indicator.

- Students were not encouraged to generate conjectures. Seventy percent of the lessons received a 1 or a 2 for this indicator.

- Appropriate connections were not made to other areas of mathematics, other disciplines, or real-life contexts. Sixty-six percent of the observers gave the lessons a 1 or a 2 for this indicator.

On the other hand, the following positive outcomes were noted for the two key indicators with the highest ratings.

- Teachers satisfactorily displayed an understanding of mathematics concepts. Eighty-one percent of the observers gave them ratings of 3 or 4, and 14% gave a rating of 5.

- Teachers respected students' prior knowledge in these lessons. Sixty-five percent of the observers gave them ratings of 3 or 4 with a mean of 2.95 (std. = .95).

The third highest rated indicator, #5, was in the area of: "the interactions reflected collaborative working relationships among students and between teacher and students." Its mean of 2.76 was deemed inconclusive, since 57% of observers gave ratings of 1 or 2.

Three additional questions in this section of the COP asked observers to rate the "likely effect" of the observed lesson in three areas, using the same Likert of 1 ("not at all") to 5 ("to a great extent"). These ratings were generally low: for "students' understanding of mathematics concepts as a dynamic body of knowledge generated and enriched by investigation," 68% received ratings of 1 or 2; for "students' understanding of important mathematics concepts," 55% of the ratings were 1 or 2; and, for "students' capacity to carry out their own inquiry," 68% received ratings of 1 or 2.

Capsule Description of the Overall Quality of the Observations — Finally, the sixth and last section of the COP instrument assessed the overall effectiveness of the observed lesson, using the same scale as for the lesson plan. Using ratings ranging from 1 (ineffective instruction) to 8 (exemplary instruction), the observers summarized their overall assessments of the quality of instruction and its likely impact of the lesson on students. Most participants (76%) were rated at least a 4, or the "beginning stages of effective instruction," with an average score of 4.24.

Summary of Year I Observations

In summary, results of the first-year data were mixed. We concluded that, while most first-year teachers who participated in this study displayed an adequate understanding of the concepts they were teaching, they still reported using far fewer standards-based strategies in their mathematics classes than they had intended to and had lesson plans that were uniformly rated lower than their classroom observations. They seemed generally successful in the areas of maintaining high levels of student engagement, working with students' prior knowledge and misconceptions, and establishing positive, collaborative norms in the classroom. A sizeable majority of the group—76%—was rated overall at about where we would hope beginning teachers would be, namely at least at Level 4 ("beginning stages of effective instruction"). Teachers received the lowest ratings in the areas of promoting students' metacognitive abilities and helping them generate their own conjectures, solve new problems, and create new knowledge.

Year II and Year I Comparisons

During 2003-2004, eleven of the above twenty-one teachers were again observed teaching a mathematics lesson. Nine stayed in the same schools as in 2002-2003, and all were observed by the same faculty member who had observed them during their first year. Eight were teaching the same grade as in the previous year. Tables 3 and 4 compare rating means for 2002-03 (Year I) and for 2003-04 (Year II) for the Lesson Plans and Observations, respectively, for several of the indicators discussed above.

For the lesson plan indicators (Table 3), the results were mixed as they were divided evenly between those that rose from Year I to Year II and those that either stayed the same or decreased. For the classroom observation indicators however, Table 4 shows that there was an increase in the presence of standards-based characteristics in the classroom. Except for one indicator (#4), all mean scores were higher in Year II than in Year I, leading us to conclude that the continuing teachers had indeed improved in their ability to facilitate students' understanding of mathematics concepts after their first year in the classroom.

Table 4 shows that observers also saw an increase in the use of activities that could be designated as standards based. The mean of the "Type of Instructional Activities Scale" rose from 1.99 to 2.17 (maximum possible rating: 3). The largest increase, however, was obtained with the rating of the "Overall Student Engagement" indicator, which rose to 2.61 from 2.40.

On the other hand, the types of cognitive activities in which students were engaged remained largely in the lowest categories of "receipt of knowledge" and "knowledge representation," despite the teachers' pre-service preparation, which had explicitly stressed inquiry-based teaching and learning, and their own stated intentions, as indicated in their surveys. (The average rating in this area did rise somewhat in Year II, to 2.15 from 1.69, out of a possible 4.) So while we conclude that their ability to deliver standards-based mathematics instruction was progressing, there was still room for improvement.

Table 3
Lesson Plan Ratings, Year I and Year Two Comparisons

Measure		**Possible Range**	**Year I** (n=21) **Mean**	**Year II** (n=11) **Mean**
	Lesson Plans			
	I. Presence of Standards-Based Characteristics			
Key 1	The lesson engages students in in-depth, higher order thinking about the content	1 – 4	2.10	2.27
Key 2	Students make real-world connections	1 – 4	2.00	2.27
Key 3	Students work together as learning community	1 – 4	2.29	2.09
Key 4	Students feel as though they have their needs met	1 – 4	2.05	2.05

Key 5	Students learn important concepts	1 – 4	2.16	2.36
Key 6	The lesson plan demonstrates effective questioning strategies for constructing student knowledge and differentiating instruction	1 – 4	2.33	2.27
II. Capsule Rating		1 – 8	2.86	3.36

Table 4
Classroom Observation Ratings, Year I vs. Year II

Measure		Possible Range	Year I Mean (n=21)	Year II Mean (n=11)
Classroom Observations				
Classroom Instructional Activity		1 – 3	1.99	2.17
Overall Student Engagement		1 – 3	2.40	2.61
Overall Cognitive Activity		1 – 4	1.69	2.12
A. Presence of Standards-Based Characteristics				
K1	Students seek & evaluate alternative modes of investigation....	1 – 5	2.38	2.77
K2	Elements of abstraction....	1 – 5	2.52	2.62
K3	Students were reflective about own thinking...	1 – 5	2.05	2.69
K4	Respect prior knowledge	1 – 5	2.95	2.83
K5	Collaborative interactions	1 – 5	2.76	3.08
K6	Promote conceptual understanding	1 – 5	2.21	2.67
K7	Students encouraged to generate conjectures	1 – 5	2.10	2.69
K8	Teacher's understanding of math...	1 – 5	3.62	3.85
K9	Connections to other areas of math, real world, etc.	1 – 5	2.11	2.42
Observation Capsule Rating		1 – 8	4.24	5.09
B. "Likely Effects" Indicators				
K10	Students' understanding of math concepts as a dynamic body of knowledge generated and enriched by investigation	1 – 5	2.32	2.55
K11	Students' understanding of important mathematics concepts	1 – 5	2.55	3.00
K12	Students' capacity to carry out their own inquiry	1 – 5	2.16	2.18
	Mean of likely effect indicators	1 – 5	2.43	2.59

Factors Contributing to Teachers' Performance—the Citywide Context

At the beginning of 2003, New York City public schools underwent a major shift as all schools were required to adopt the *Everyday Mathematics* program for grades one through five, and *Impact Mathematics* for grades six through eight [11,12]. While these programs are generally consonant with the standards-based reforms, teachers struggled to implement them. Students were caught in the gap between their previous curriculum and the new programs. Frequently, for a particular grade level, the new curriculum would be aimed at a certain level, and assumed knowledge of content to which the students had never been exposed. Teachers consistently expressed their lack of preparation and training in the new programs and struggled to understand and implement them. They reported very little support from their schools, in terms of professional development or mentoring, for making this adjustment.

Related to this issue is that of the impact of mandated assessments, sometimes those of the new mathematics programs, but also the citywide mathematics examinations. Several teachers also complained about the need to concentrate on test prep materials prior to the citywide mathematics tests in May. In the words of one:

> The biggest thing that comes back to me is the amount of test prep at the end of the year. I would try and abandon the program because it was not working. Example: I was supposed to tell the students [the times of] sunrise and sunset and have them calculate the length of the day. I had students who couldn't tell time. High and low temperature. They couldn't tell temperature.

In all, ten of the twenty-one teachers specifically mentioned difficulties resulting from a lack of time and/or flexibility to teach appropriately, due to either the switch to the new mathematics programs, the demands of test preparation for the standardized exams, or both.

Factors Contributing to Teachers' Performance—School-Level Support

Of major concern to educators is the question of the general environment of particular schools, sometimes referred to as the school culture or as the school context. Through teacher surveys and interviews, we focused on three areas of school-level support for these novice teachers as they attempted to teach math for understanding: 1) that provided by their administrations; 2) that provided by their colleagues in general; and, 3) the extent to which they were engaged in conversations about teaching and learning math.

<u>Administration</u> — Teacher surveys asked teachers about the ways in which their schools supported excellence in math instruction, and the barriers they faced in teaching math. Nine of the twenty-one teachers found their schools to be supportive; of these, four referred to helpful staff developers or other veteran teachers, three mentioned supplies of books or manipulatives, and two mentioned feedback from assistant principals following observations. Two-thirds of the respondents in 2002-2003 and nine of the eleven in 2003-2004 identified school-based barriers, other than the ones related to the new math programs and mandated assessments that were discussed above. When asked to describe these barriers, the teachers singled out: lack of classroom demonstrations and guidance; lack of resources (books, materials); and, class size, scheduling, and time constraints.

In their second year, teachers were asked to rate their school administration on several dimensions. Responses could range from 1 (strongly disagree) to 4 (strongly agree). Table 5 reports the results.

Table 5
Teacher Ratings of Administrative Support for Teachers, Inquiry (n=11)

Item	Percent Responding: 1, 2 Strongly disagree, disagree	Percent Responding: 3, 4 Agree, strongly agree	Mean
1. The principal at this school:			
a. Promotes parental and community involvement in the school....	30	70	2.8
b. Works to create a sense of community in this school.........	80	20	2.1
c. Takes a personal interest in the professional development of teachers....	50	50	2.4
d. Is strongly committed to shared decision making............	70	30	2.1
e. Ensures that student learning is the "bottom line" in the school....	78	22	2.2
f. Supports and encourages teachers to take risks.....	100	0	1.8
g. Is a strong leader in school reform.........	78	22	1.7
2. Teachers in this school are well-supplied with materials for investigative instruction.................	60	40	2.3
3. Assessment of student performance leads to changes in our school's curriculum........	54	44	2.3

As seen here, teachers generally found their principals unsupportive. None reported that their principal encouraged them to take risks; only 22% agreed that their principal ensured that student learning was the "bottom line" at the school, etc.

Colleagues — A second series of questions on the teacher survey in Year II asked teachers to measure the extent to which a shared culture of learning existed among the school faculty. Responses could range from 1, (strongly disagree) to 4 (strongly agree). Table 6 presents the results.

Table 6
Teacher Ratings of Colleagues: Shared Values, Behaviors, and Support (n=11)

Item	Percent Responding: 1, 2 Strongly disagree, disagree	3, 4 Agree, strongly agree	Mean
Shared values:			
It is important for students to have input in establishing criteria by which their work will be assessed............	30	70	2.8
Teachers in this school have a shared vision of instruction......	56	44	2.2
Our stance towards our work is one of inquiry and reflection.....	50	50	2.8
Behaviors:			
Teachers are engaged in systematic analysis of student performance data......	44	56	2.7
Teachers in this school are continuously learning and seeking new ideas....	20	80	3.1
Teachers in this school are well informed about the national, state, and professional education standards, e.g., AAAS, NRC, NCTM, for the grade levels they teach....................	25	75	2.9
Support:			
I feel supported by colleagues to try out new ideas......	10	90	3.1
Teachers in this school trust each other......	56	44	2.4
Teachers in this school feel responsible to help each other do their best...	20	80	2.6

As seen above, teachers gave their colleagues "middling" ratings on most of the questions. While they generally found their colleagues supportive, less than half felt that their school had a shared vision of teaching and learning, and only half felt that "*our stance is* one of inquiry and reflection."

Conversations About Teaching Mathematics — In the interviews conducted in 2003-2004, when discussing individual students, teachers were asked, "Did you ever discuss teaching math to this student with another teacher, a supervisor, or a math coach or mentor?" Each of the six teachers interviewed described three students in depth during the interviews; hence, this question was asked about eighteen students. Teachers reported discussing twelve of the eighteen. However, the topic of the conversation was often one of general concern, such as getting a better class placement, or about areas other than math, e.g., reading difficulties. Only four of the twelve conversations reported had to do specifically with some aspect of the student's performance in math. The interviews give the impression that, in general, teachers discussed their students with others in the school most often when there was a problem, sometimes out of general interest and concern (e.g., checking with the current teacher of a former student to see how s/he was doing), but only rarely around specific issues related to math teaching and learning.

Summary and Conclusions

This study tracked the experiences of a cohort of novice teachers as they moved from their college classes into public elementary schools in New York City. As noted, there was a sharp decline between their pre-service intentions and their reported classroom practices in the first year. Overall, observers gave the first-year teachers mixed ratings in terms of their implementation of standards-based classroom practices. The strongest points of the novice teachers were their ability to engage students and their understanding of the content they were teaching. However, they still lacked proficiency in guiding students to higher levels of conceptual understanding by leading them to reason, examine their own thinking, and create their own knowledge.

As the teachers moved into their second year in the classroom, their abilities to teach math for conceptual understanding increased, reaching Level 5, or "solid beginning stages of effective instruction" on the capsule rating, on average. However, their average ratings on most key indicators still hovered around the mid-point of 2.5 on the scale from 1 to 5; we would have hoped to see averages of at least 3 by the second year. Their strongest ratings remained in the areas of their own content knowledge and in their ability to create a collaborative atmosphere in the classroom.

Such findings are in alignment with other recent studies of the impact of standards-based preparation of mathematics teachers. Weiss reports that even teachers who were provided with extensive in-service professional development to enhance their ability to use standards-based materials via NSF's Local Systemic Initiatives, "struggled with key elements of teaching for understanding" [13].

The school context factors listed below seem to have contributed to the ability to teach math for understanding of the new teachers in our study.

People Resources, More than Material Ones — For mathematics instruction at least, new teachers singled out the role of mentoring and coaching, and specifically requested more demonstration math lessons. Although they did mention the need for more manipulatives, computers, or other material resources, these were clearly less of an issue than was the need for ongoing professional support and development.

Difficulties Arising from the Citywide Adoption of New Math Programs — Specifically, it is quite clear that teachers would have benefited from more training in using the new math programs in 2003-2004. Many preferred more flexibility in choosing teaching materials and texts.

School Culture — School leaders generally got very low marks in terms of focusing on student learning, leading for change, and for supporting shared decision making and risk taking on the part of their staffs. While teachers generally reported higher levels of support from fellow teachers in their schools, their responses still indicate considerable room for improvement in the cultures of these schools, in terms of promoting systematic teacher inquiry and reflection about teaching and learning.

Limitations of the Study

Although the results of this study are in keeping with the results of other recent ones, caution needs to be exercised in several areas. The first concerns whether the sample was truly representative. It is probable that the most confident and able teachers self-selected themselves for the additional scrutiny that this study entailed. The second caution concerns whether only one observation each year gave a valid indication of the teachers' capabilities. Anecdotally, the designers of the CETP core evaluation study design report that in other studies increasing the number of observations per year to two did not change the ratings significantly. Additionally, we suspect that two conflicting tendencies may cancel each other out in this area. The first is the extra preparation and planning that went into these lessons because they were being observed which would have resulted in an overestimation of their daily teaching practice. On the other hand, not even the most expert teacher in the world could demonstrate his/her entire repertory of standards-based teaching strategies in one forty-minute lesson; hence, this study may have underestimated their abilities. Taken together, these two tendencies may have effectively counteracted each other.

Implications for Pre-Service and In-Service Professional Development

What circumstances are necessary, in order for these and other novice teachers to teach mathematics for understanding at the levels that they aspired to as undergraduates? Despite their training and strong belief in using standards-based practices, the level of cognitive activities engaged in by their students seldom reached the highest levels of knowledge construction and knowledge representation. Some of the obstacles the teachers faced were school based: class size, scheduling and time constraints, lack of materials and resources, and lack of sustained training and guidance in using the new mathematics programs. These were compounded by the pressure of state-mandated, high-stakes tests. However, we may also ask to what extent their pre-service education classes had adequately prepared them for and assessed their abilities to apply theoretical understandings of standards-based pedagogy in practice. Teaching for conceptual understanding not only requires addressing the school-based obstacles mentioned above, it also requires a high level of teacher skill in selecting and designing activities. As other studies previously found, beginning teachers tend to move away from their beliefs and "give up their new conceptions as they struggle to survive and to fit into the instructional norms of traditional educational practices" [14]. The consequences of this struggle could be devastating to the students they teach, especially those in urban areas who already face tremendous challenges.

Certain recommendations are therefore obvious. From both their pre-service and their in-service training, teachers need more scaffolding, more model lessons, and more specific support for engaging students at higher cognitive levels, for helping students to create their own understandings, and for implementing cooperative learning. School administrators need to be convinced that teaching mathematics for understanding cannot be accomplished by teaching to the test. Curricular materials such as *Everyday Mathematics* provide teachers with standards-based instructional activities [11]. However, teachers must be empowered to deviate from these activities in order to address their students' needs when necessary. Finally, schools need to initiate conversations with teachers about not only teaching and learning in general, but the specific progress of individual students, and to sustain those conversations in an ongoing way throughout the school year. ■

Acknowledgment

The research upon which this article is based was supported by a grant from the National Science Foundation to the Center for Advanced Study in Education, CUNY.

References

[1] *Principles and Standards for School Mathematics*, National Council of Teachers of Mathematics, Reston, VA, 2000.

[2] L. Ma, *Knowing and Teaching Elementary Mathematics: Teachers' Understanding of Fundamental Mathematics in China and the United States*, Lawrence Erlbaum Associates, Mahwah, NJ, 1999.

[3] M.A. Timmerman, "The Influences of Three Interventions on Prospective Elementary Teachers' Beliefs about the Knowledge Base Needed for Teaching Mathematics," *School Science & Mathematics*, **104**(8) (2004) 369-382.

[4] M.J. Behr, G. Harel, T. Post, and R. Lesh, "Rational Number, Ratio, and Proportion," in D. A. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning*, Macmillan Publishing Co., NY, 1992.

[5] D.A. Grouws and K.A. Schultz, "Mathematics Teacher Education," in J. Sikula, T.J. Buttery, and E. Guyton (eds.), *Handbook of Research on Teacher Education*, Macmillan Publishing Co., New York, 1996.

[6] M. Lampert, "When the Problem Is Not the Question and the Solution Is Not the Answer: Mathematical Knowing and Teaching," *American Educational Research Journal*, **40** (1990).

[7] D. Ball, H. Hill, and H. Bass, "Knowing Mathematics for Teaching," *American Educator*, **44** (2005) 14-16.

[8] U.E. Wagener, "Changing the Culture of Teaching," *Change*, **23**(4) (1991) 28-37.

[9] F. Lawrenz, D. Huffman, and K. Appledoorn, *Classroom Observation Handbook.* CAREI, College of Education & Human Development, University of Minnesota, Minneapolis, MN, 2002.

[10] *Professional Standards for Teaching Mathematics*, National Council of Teachers of Mathematics, in D.F. Steele, "The Interfacing of Pre-Service and In-Service Experiences of Reform-Based Teaching: A Longitudinal Study," *Journal of Mathematics Teacher Education*, **4** (2001) 169.

[11] *Everyday Mathematics*, University of Chicago School Mathematics Project, McGraw-Hill, New York, NY, 2004.

[12] *Impact Mathematics*, McGraw-Hill, Glencoe/McGraw-Hill, 2002.

[13] I.R. Weiss, "Evaluation of the NSF Local Systemic Change Program: Implications for Preparing New Teachers of Mathematics," *Association of Mathematics Teacher Educators (AMTE) Connections*, **15**(1) (2005) 12.

[14] S. Wilcox, P. Lanier, P. Schram, and G. Lappan, *Influencing Beginning Teachers' Practice in Mathematics Education: Confronting Constraints of Knowledge, Beliefs, and Context*, The National Center for Research on Teacher Learning, Michigan State University, East Lansing, MI, 1992.