Experimental Methods and Results in a Study of PBS TeacherLine Math Courses

By
Paula S. Dominguez, Ed.D.
Craig Nicholls, Ph.D.
Barbara Storandt, M.Ed., M.S.
Hezel Associates, LLC


Paula Szulc Dominguez, Ed.D.


Craig Nicholls, Ph.D.


Barbara Storandt, M.Ed., M.S.

Introduction

Online teacher professional development has received considerable attention for its ability to provide teachers with access to outstanding resources in a flexible, ongoing manner. Its timing is impeccable. The requirements of the No Child Left Behind Act have thrown the issues of teacher capacity and effectiveness into the spotlight. NCLB's provisions call for teachers to develop and employ a heightened ability to respond to feedback from standardized testing in a manner that meets the demands of external accountability (Halverson et al., 2005). In addition to diagnosing and reacting to feedback from tests and other data, leading today's classrooms require teachers to accommodate the needs of diverse students who learn in different ways and to respond dynamically to unforeseen learning opportunities (Darling-Hammond et al., 2005). All of this work matters, because the interaction which occurs between teachers and students in the classroom exerts considerable impact on student learning (Wenglinsky, 2002).

To have teachers in place who can respond to these competing demands, schools and school districts must provide them with professional development opportunities that exhibit qualities associated with "high quality" approaches to teacher learning, such as a focus on content knowledge, opportunities for active learning, and coherence with other learning activities (Garet et al., 2001). Against this backdrop, providers of online education have stepped up to offer professional development models that can also demonstrate "high quality" standards. Done well, online approaches can offer teachers valuable experiences that are unavailable through traditional venues. For example, as Thomas (2004) points out, online professional development provides access to content and instruction over time, with the potential for participants to continue their dialogue before and after the course has been completed. In this manner, online approaches overcome the "one time contact" disadvantage associated with some traditionally delivered professional development courses.

PBS sought to build on the strengths of high quality professional development through an online venture funded under the U.S. Department of Education's Ready To Teach (RTT) program. PBS TeacherLine was designed to provide high-quality, facilitated online professional development for K-12 teachers nationwide. Through PBS TeacherLine, more than 20,000 educators have participated in 90 online, facilitated courses in reading, mathematics, science, instructional strategies, instructional technology, and curriculum mapping.

To what extent has PBS TeacherLine been successful in creating an online model of teacher professional development that gets to the heart of education needs – impacting teacher beliefs, classroom instruction, and student learning? This paper presents the results from one component of the comprehensive external evaluation of PBS TeacherLine: an experimental study that unfolded in the fifth and final year of the RTT grant. The external evaluation was conducted by the Hezel Alliance, a partnership composed of Hezel Associates, an independent consulting firm, the Education Alliance at Brown University, and Syrtis.

Background

The first two years of PBS' RTT grant centered on the development of a model for online professional development, facilitator recruitment and training, and course creation. In addition, PBS solidified its distribution channel by supporting partnerships comprised of local stations working with local education agencies. Some online courses began to be offered in Year 2, but it was in Year 3 of the RTT grant that a more widespread array of PBS TeacherLine courses became available to stations nationwide.

The mixed methods evaluation of PBS TeacherLine included the collection and analysis of data information regarding course quality, course participation and completion, participant satisfaction, and quality of online interaction. In Year 4, the U.S. Department of Education presented the evaluation team with the challenge of incorporating "gold standard" experimental research into the evaluation. In response, the evaluation team designed and executed a randomized experimental study that was organized and carried out over a two-year period.

Methods

The experimental study addressed the question of whether and to what extent PBS TeacherLine participation impacted teachers' attitudes, instructional practices, and student achievement. The design assumed a logical chain of events: that is, it assumed that change in teachers' attitudes precedes change in teachers' instructional practice, which in turn precedes improvements in students' academic performance. The evaluation team developed or selected measures to assess each link in this chain.

Participating in the experiment were a (final) total of 92 elementary teachers in grades three to five, who were recruited from public schools in three locations: Miami-Dade County (Florida), Richland Two (South Carolina), and Buffalo (New York). The teachers were assigned at random into one of two groups. Teachers in the treatment group participated in two online PBS TeacherLine courses related to elementary mathematics during the 2004-05 school year (Patterns and Relations: Algebra Concepts for Grades 1-5," and "Shaping Up: Teaching Geometry Using Technology in Grades 3–5) , while members of the control group did not participate in any online courses.[1] All participants received an honorarium for their time and efforts.

At the start of the study and at its completion, all teachers completed an online survey regarding their attitude toward mathematics teaching and learning. The attitude measures assessed both teachers' attitudes toward reformed mathematics teaching and learning, and their attitudes toward traditional instructional practices. Some items were adapted from Horizon Research's 2000 National Survey of Science and Mathematics Education. Others were developed based on dimensions assessed by the Reformed Teaching Observation Protocol, the observation protocol the evaluation team employed in its site visits. The survey was worded to reflect the instructional value teachers attach to various teaching and learning activities, rather than the frequency with which they employed these activities.

Originally, separate measures were drafted for teaching activities and learning activities. Within each dimension, items were written to reflect both reformed and traditional practices. The expectation was that the value teachers attached to these activities would be inversely related to each other, in that the more teachers valued reformed teaching and learning practices the less they would value traditional practices. When the baseline survey data were analyzed, however, it became clear that this assumption was not upheld, and also that there was no strong distinction between teaching and

[1] Note that teachers in the control group were not prevented from participating in any other professional development experiences during the course of the study.

learning activities as seen by the teachers. Therefore, questions on reformed teaching and learning activities were grouped together on one scale, and questions on traditional teaching and learning activities were grouped together on another. Internal consistency as measured by Cronbach's alpha was .87 for both scales at baseline and slightly higher for both at posttest (.88 and .89). A total of 84 teachers completed both pre and post surveys.

To track teachers' instructional practice, the evaluation team selected the Reformed Teaching Observation Protocol (RTOP). The RTOP was designed, piloted, and validated by the Evaluation Facilitation Group of the Arizona Collaborative for Excellence in the Preparation of Teachers for use evaluating science and mathematics instruction in grades K-20. It measures central characteristics of reformed mathematics and science instruction as propounded by the standards developed by the major professional organizations such as NCTM.

The RTOP requires a detailed narrative of the mathematics lesson with an accompanying time log, a description of the physical characteristics of the classroom with an emphasis on mathematics and technology materials, and a set of 25 items, each to be rated on a scale of 0 (no evidence) to 4 (highly descriptive of the lesson), resulting in a possible total score range of between 0 and 100. The RTOP's rating section is divided into five subscales, each focusing on elements of reformed mathematics practice.

The classroom observation data were collected by 11 members of the evaluation team. Three expert raters conducted training for the remaining eight data collectors. Training for data collectors occurred through intensive full-day trainings with homework assignments, followed by joint observations, consultative coding sessions, and follow-up consultation through data review. The training focused on developing a shared understanding of the RTOP's key constructs and establishing an evidence-based rationale for rating the observed lessons across the data collection team. After the full-day trainings, expert raters accompanied the data collectors to a classroom observation. The jointly observed lesson was independently rated by each observer, with the expert providing feedback and clarification as needed. Subsequently, the expert raters provided on-going consultation to the data collectors as particular questions arose.

All RTOP narratives and ratings submitted by the data collectors were reviewed by a team of three expert raters to ensure consistency, quality, and comprehensiveness within and across observers. This review process was designed to address the potential for individual coder drift and to monitor the overall level of inter-coder agreement (Taplin & Reid, 1973). Expert reviewers changed any discrepant ratings, noting the original score, the changed score, and the rationale for the change on each RTOP file. Expert reviewers met regularly to discuss and resolve any ambiguities in evidence or ratings.

Data collectors observed a total of 288 mathematics lessons in the three districts, and completed RTOP narratives and ratings for all of them (although not all of these observations were included in the final data set; e.g., some were considered to be incomplete). The observations took place at three times: at the beginning of the study, at the time that corresponded to the completion of treatment group member's first online course, and at the end of the school year.

Finally, to document student learning the evaluation team relied on two, 20-item student tests, one each for Algebra and Geometry, developed by university specialists in mathematics education as pre- and post-tests for this study for each of the three grade levels. The tests were developed to align with the NCTM *Standards* and piloted in elementary classrooms. Individual student responses were matched using a series of demographic questions that appeared on both pre- and post-tests. Student

birthday and students' phone number (last 4 digits) were the primary demographics used for matching student records; however, the evaluation team also looked at student gender and ethnicity when missing data existed. In total ,this process resulted in 1,137 usable student records; 339 from students in the Miami-Dade County School District (FL), 322 from students in the Buffalo City School District (NY), and 476 from students in Richland County School District Two (SC).

In addition, the evaluation team received permission from two of the school districts (Miami-Dade and Richland Two) to access students' state standardized test data in mathematics for the current and prior years. South Carolina's statewide assessment program to measure student performance on state standards is the Palmetto Achievement Challenge Test (PACT). The PACT is administered to all students in grades three through eight each year. PACT data from 516 fourth and fifth grade students in the Richland Two district (SC) for 2004 and 2005 were analyzed (since third grade students did not take the 2004 PACT test, they were excluded from the analysis). To analyze the PACT data, the evaluation team calculated gain scores based on the difference between numeric proficiency scores.

Findings

As described above, the experimental research examined a chain of events – from change in teachers' attitudes to change in teachers' instructional practice to improvement in student academic performance. Findings from the assessments of each of these links are presented below in this same sequence, beginning with baseline findings and concluding with the results of statistical tests on teacher or student gains.

Teachers' attitudes toward reformed math teaching and learning

Across the full sample at baseline, teachers indicated a stronger belief in the value of reformed math practices than in the value of traditional practices (see Table 1; note that the former scale contains one more item, which could raise its mean by up to 4 points). No differences were seen between treatment and control groups in baseline attitudes on either scale.

Table 1. Teacher attitudes at baseline

Scale	N	Min	Max	Mean	SD
Traditional scale					
Treatment	44	12	35	25.32	5.246
Control	43	17	37	26.05	4.815
Total	**87**	**12**	**37**	**25.68**	**5.022**
Reformed scale					
Treatment	44	28	44	38.59	4.222
Control	43	25	44	38.67	4.592
Total	**87**	**25**	**44**	**38.63**	**4.383**

Attitudes toward reformed math teaching and learning practices were uncorrelated with attitudes toward traditional practices ($r = -.01$, $p = .93$). This suggests that teachers can believe in reformed practices (e.g., greater student involvement, initiative, group work) and at the same time believe that traditional, independent desk work and teacher-centered instruction is valuable. This was unexpected—it was thought that the two would be negatively correlated.

Belief in reformed teaching and learning practices at baseline was weakly to moderately correlated with baseline RTOP scores across the full sample ($r = .38$, $p < .001$). Baseline RTOP was uncorrelated with belief in traditional teaching and learning practices, providing evidence for the construct validity of the scales, in that the instructional practices teachers say they value are reflected in their practice.

Table 2 shows gains in the Reformed Math scale over the course of the experiment. Members of the treatment group demonstrated changed scores with regard to their belief in reformed math teaching and learning principles over the course of the year relative to participants in the control group. The difference, less than two points, is small but statistically significant. The test does not assume equality of variance in mean gain between the groups.

Table 2. Change in attitude toward reformed math teaching and learning principles

Year	N	Mean gain	SD	t	Df	Sig.
Treatment	43	2.12	2.15	2.31	60.7	.024
Control	41	.49	4.00			

No other variables were significant predictors of change in attitude. The following variables were tested along with treatment group:

Years taught at the K-12 level	Gender	Race/ethnicity
NCTM familiarity	Education level	Grade level

There was also no change overall or by group in teacher attitudes toward traditional forms of teaching and learning. These findings suggest that the sequence of PBS TeacherLine math courses had a real, if modest, effect on changing teacher attitudes toward teaching and learning principles that are widely understood to be educationally valuable and embodied as such in NCTM standards.

Classroom teaching practices

Average scores for reformed math teaching practices were low across the whole sample at baseline. No significant differences existed (at baseline) between treatment and control groups, or between grade levels. However, NY teachers started out significantly lower on RTOP ratings than the other districts, and women significantly outperformed men. As noted above, teachers who scored higher on the baseline attitude survey also tended to score higher on the baseline RTOP. Table 3 provides a breakdown of baseline RTOP scores by state/district and treatment group.

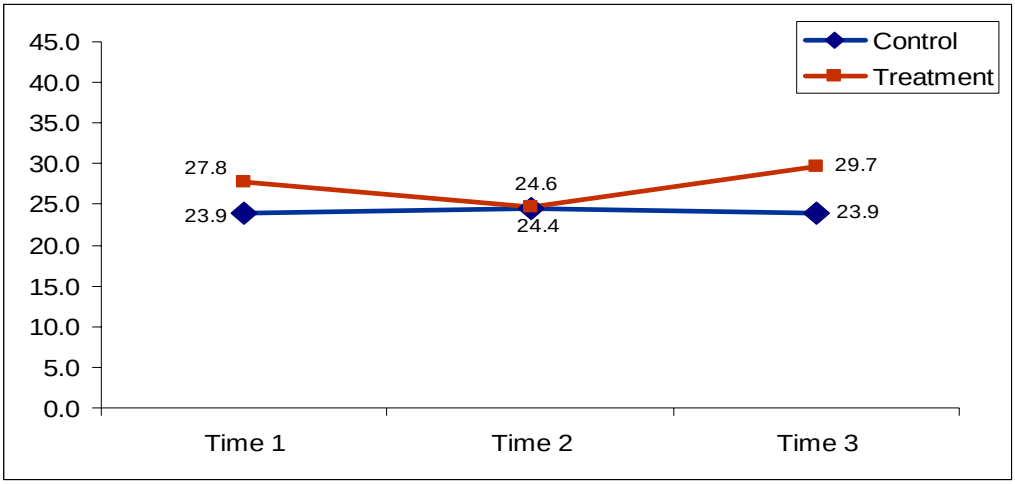
Hierarchical Linear Modeling (HLM) was used to model within-teacher growth over time in RTOP ratings as a function of teacher characteristics including treatment condition. Teachers' education level, familiarity with NCTM standards, baseline attitudes, gender, race/ethnicity, grade level, and district were all considered in this analysis in addition to treatment condition.

Table 3. Baseline RTOP scores

State	Cohort	N	Min	Max	Mean	S D
FL	Control	15	10	45	25.00	10.63
	Treatment	13	10	47	28.85	10.78
	Total	**28**	**10**	**47**	**26.79**	**10.68**
NY	Control	15	5	52	22.67	14.85
	Treatment	12	8	42	20.00	9.22
	Total	**27**	**5**	**52**	**21.48**	**12.51**
SC	Control	15	11	43	24.93	8.64
	Treatment	20	9	54	31.40	14.48
	Total	**35**	**9**	**54**	**28.63**	**12.59**
Total	Control	45	5	52	24.20	11.45
	Treatment	45	8	54	27.62	12.89
	Total	**90**	**5**	**54**	**25.91**	**12.24**

Although teachers in the treatment group appeared to move toward more reformed math teaching practices as measured by the RTOP, compared to the control group, this trend is not statistically significant. Figure 1 shows the unadjusted change in mean RTOP scores from Time 1 to Time 3. Controlling for the other variables listed above did not change the results for the effect of the treatment.

Figure 1. Trend over time in RTOP ratings



Variables found to affect change in RTOP ratings over time (across the full sample) were district and education level. Regardless of treatment group, NY teachers improved their RTOP ratings significantly compared to FL and SC teachers. Also, better educated teachers tended to show greater movement toward reformed practices (although this was of borderline significance). [2] Table 4 shows the estimated effects of the variables influencing baseline RTOP and change in RTOP score over time.

Table 4. Estimated Effects on RTOP scores

Fixed Effect	Coefficient	Standard Error	t	df	p
For INTRCPT1, P0					
INTRCPT2, B00	28.077177	1.389769	20.203	83	0.000
PSUMA2B2, B01	0.814281	0.239977	3.393	83	0.001
GEND, B02	-6.979259	3.067235	-2.275	83	0.025
NYVSFLSC, B03	-5.997403	2.451282	-2.447	83	0.017
For TIME slope, P1					
INTRCPT2, B10	-0.489222	0.872747	-0.561	84	0.576
NYVSFLSC, B11	3.323795	1.572179	2.114	84	0.037
EDUC, B12	1.124496	0.684331	1.643	84	0.104

Table 4 shows that each one-point swing (positive or negative) in baseline attitude (PSUMA2B2) was associated with a .81 point movement in the same direction in baseline RTOP. Men scored about 7 points lower than women, and NY teachers averaged about 6 points lower than SC and FL teachers. Also, NY teachers improved about 3.3 points more than FL or SC teachers at each successive

[2] The best-fitting model, which nonetheless left considerable unexplained variance in both the intercept (baseline mean) and time slope is as follows:

Level-1 Model

$$Y = P0 + P1*(TIME) + E$$

Level-2 Model

$$P0 = B00 + B01*(PSUMA2B2) + B02*(GEND) + B03*(NYVSFLSC) + R0$$
$$P1 = B10 + B11*(NYVSFLSC) + B12*(EDUC) + R1$$

where the outcome Y is RTOP score, PSUMA2B2 is the score for teachers' baseline belief in reformed teaching and learning practices, GEND is teacher gender (coded 0 for women and 1 for men), and NYVSFLSC is a dichotomous variable differentiating New York from Florida and South Carolina (coded 1 for NY, and 0 for FL and SC).

observation point, and each point above or below the mean on the 4-point Education Level scale influenced change in practice (as measured by the RTOP) by a little over a point.

In summary, participation in the two PBS TeacherLine courses we examined was not associated with a significant movement toward reformed teaching practices as measured by the RTOP. This may mean either that practice did not improve, or that some improvement occurred but was not captured by the observation protocol. Reformed teaching practices were influenced by other variables, however, including school district, gender, education level, and attitudes.

Student academic performance

Table 5 shows students' pre-test, post-test, and gain scores in Algebra, and Table 6 shows the same in Geometry. There were no significant differences between groups at pre-test on either of the measures. Across the sample, gains from pre- to post-test were modest for both the Algebra and Geometry tests. HLM was used to test the effect of the treatment on student gain scores while accounting for student and teacher level variables.

Table 5. Mean Algebra Scores and Gain

	N	Min	Max	Mean	SD
Pre-test					
Control	534	0	20	9.90	3.690
Treatment	600	1	20	10.31	3.963
Total	1134	0	20	10.12	3.841
Post-test					
Control	534	0	20	11.99	4.103
Treatment	600	1	20	12.34	4.434
Total	1134	0	20	12.17	4.283
Gain (Post – Pre)					
Control	532	-15	13	2.10	3.590
Treatment	599	-7	11	2.02	3.211
Total	1131	-15	13	2.06	3.393

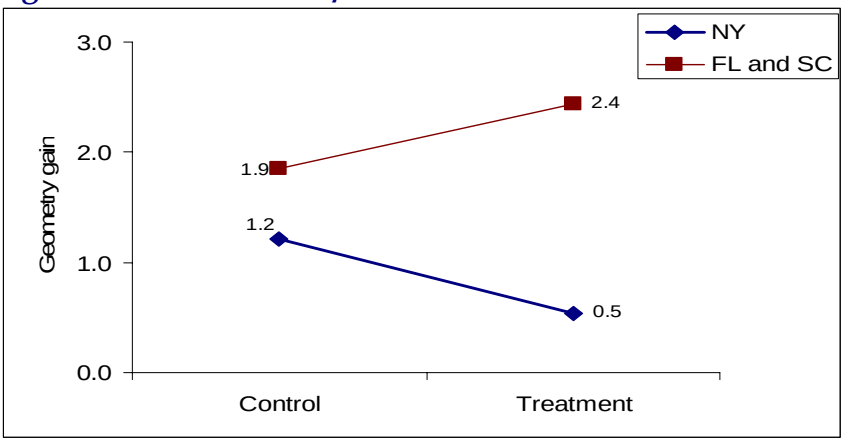
Table 6. Mean Geometry Scores and Gain

	N	Min	Max	Mean	SD
Pre-test					
Control	527	0	20	10.37	3.279
Treatment	577	1	19	10.63	3.260
Total	1104	0	20	10.51	3.270
Post-test					
Control	523	0	20	12.14	3.807
Treatment	596	0	20	12.52	3.781
Total	1119	0	20	12.35	3.797
Gain (Post – Pre)					
Control	515	-11	11	1.71	3.501
Treatment	573	-13	15	1.98	3.473
Total	1088	-13	15	1.85	3.487

No significant difference was found between treatment and control groups on gains for either of the tests, and the student variables of gender and ethnicity did not influence gains. The inclusion of teacher-level variables (attitudes, education level, NCTM familiarity, race/ethnicity, and grade level) does not change the findings. However, students of FL and SC teachers significantly outperform students of NY teachers in terms of gains in both Algebra and Geometry. There is also a nearly significant interaction between treatment and district, in which NY students of control group teachers outperform students in the treatment group, while the reverse holds true for SC and FL (Figure 2).[3]

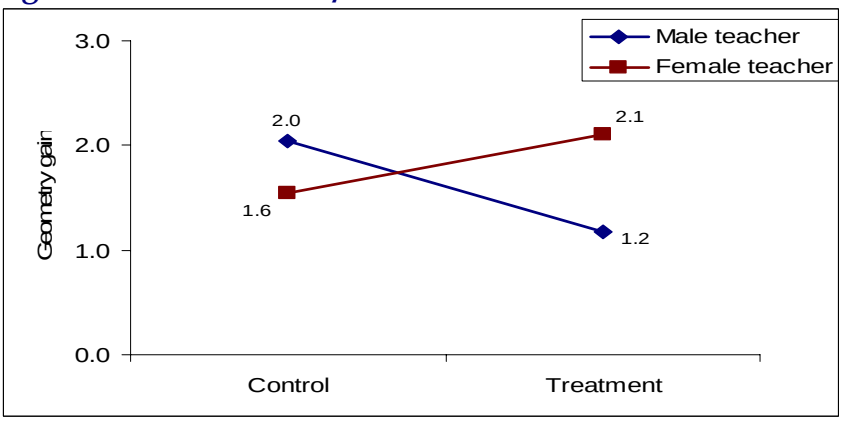
[3] Running the analysis only within South Carolina and Florida did not lead to significant results for the treatment group.

Figure 2. Treatment/District interaction in student Geometry gain



There is a nearly significant interaction between treatment group and teacher gender in their effects on student gains in Geometry (Figure 3).

Figure 3. Treatment/Gender interaction in student Geometry gain



However, the number of male teachers in this sample was small and it is unclear if the sample is representative of male teachers in general. In addition, removing them from the analysis did not result in statistically significant gains among students of the female treatment teachers. Finally, limiting the analysis to female teachers in FL and SC did not result in a significant advantage for treatment teachers.

Data from the Spring 2005 Palmetto Achievement Challenge Test (PACT) (South Carolina's state assessment), as well as from Spring 2004 for the same students, were obtained both as scale scores and as proficiency levels. Only the 4th and 5th grade students in our sample took the PACT in the prior as well as the present year, so the analysis is limited to them. Using HLM, tests were conducted on the Spring 2005 scale scores, controlling for Spring 2004 scores, and separately on a constructed variable representing change in proficiency level from Spring 2004 to Spring 2005. As with the other tests of student academic performance, no significant effect of the treatment was found in either analysis, nor did the student variables of gender and ethnicity show an influence on gains.

Discussion

According to our results, although the treatment group's ratings of their beliefs in reformed math teaching and learning principles changed over the course of the year, relative to those of the

comparison group, PBS TeacherLine participation did not appear to influence teacher practice, as measured by the RTOP. Students' outcomes were similarly unaffected by their teachers' participation in online teacher professional development courses. These findings may indicate that practice and student learning did not change, or that some change might have occurred that was not captured by the instruments we employed.

The first implication of these findings is that they appear to support the hypothesis that changes in teachers' beliefs about instruction may precede observable changes in their practice. This directional model of teacher professional development is not universally upheld by other researchers (most notably Guskey, e.g., 1986), who suggests that student learning outcomes inform changes in teacher practice).

A second implication from our work is that teacher practice is robust. The two PBS TeacherLine courses involved (officially) a total of 60 hours of work (from other reports, it is more likely that the two courses called for between 80 to 100 hours of teachers' time). Despite this, the teachers in the treatment group did not demonstrate any observed differences in their practice at the conclusion of the study.

A final implication of the study is that there are large contextual differences across the participating districts that appear to contribute meaningfully to the findings. Specifically, reformed teaching practices were influenced by school district (as well as gender, education level, and attitudes), while student gains in both Algebra and Geometry were associated with their school district location. We did not "unpack" school district contextual variables (such as textbooks used, other school-wide or district-wide professional development efforts underway), but clearly, such information is vital to understanding at a deeper level the ways that locale might interact with the impact of a professional development activity.

While findings from this experiment are important, the evaluation team acknowledges several important limitations. Our study tested a defined grade range, only two online courses, and a particular subject area – mathematics – with a distinct status within the elementary curriculum. Other research indicates that elementary mathematics practice is difficult to change, for a number of reasons. As a whole, elementary teachers report that they feel less well prepared to teach mathematics than reading or language arts (Weiss, 1994). Elementary mathematics teachers have been shown to have a weak and uneven understanding of mathematics as a subject matter (Ball, et al., 2001), particularly as compared to teachers from other countries (Ma, 1999).[4]

It may be that changing elementary mathematics practice requires a much more substantial and long-term set of interventions and systemic changes than any one professional development experience can provide. It is our view that the finding that the PBS TeacherLine courses under consideration positively impacted the "first link" in the causal chain leading to improved student achievement represents an important first step. For teachers to change their classroom practice, it may require a longer period of time to allow them to confront and try out alternative approaches and put into place their own new models of instruction. The lesson for the evaluation team is that PBS TeacherLine content that teachers work with needs to be better targeted and the design must be structured to capture over a longer period of time the incremental growth associated with teachers' changed practice. PBS TeacherLine was recently granted another five-year Ready To Teach grant, and the evaluation team will put into place a series of experiments and a longitudinal study that build on the lessons learned and findings from the research reported here.

[4] Many thanks to our colleagues at the Education Alliance at Brown University for noting the different challenges associated with changing practice across subject areas.

About Hezel Associates

As an independent research firm with two decades of expertise in the evaluation of school-based communitywide, and federally-supported efforts to improve students' well being. Hezel Associates has worked closely with state agencies, K-12 schools, community organizations, institutions of higher education, and employers throughout the nation to satisfy increasingly higher expectations for analyzing performance targets. We understand the importance of providing state education leaders with empirical, objective evidence examining program implementation, progress, and outcomes.

Paula Szulc Dominguez, Ed.D.

Dr. Paula Szulc Dominguez has worked with Hezel Associates since 1991. Leading Hezel Associates' research and evaluation efforts, Dr. Dominguez has overseen multi-site formative, process and summative evaluations and experimental research. She has supervised studies on behalf of schools, non-profit organizations, state education agencies, institutions, and systems of higher education, and multi-partner consortia. Dr. Dominguez has authored publications and presented at national conferences on evaluating dissemination outcomes and understanding models of technology-enhanced learning.

Craig Nicholls, Ph.D.

Dr. Craig Nicholls is the Research and Evaluation Manager for Hezel Associates, overseeing daily implementation of the research and evaluation projects for clients such as PBS TeacherLine, the Concord Consortium and South Carolina's Early Review Team initiative. In this position, Dr. Nicholls continuously balances rigorous research and evaluation methods with client relationships and satisfaction.

Barbara Storandt, M.Ed., M.S.

Barbara Storandt, Research Associate for Hezel Associates, supports the development and execution of programs and evaluations of educational interventions, products, services and organizational processes. Ms. Storandt has two master's degrees in education, and she is certified in New York to teach biology for grades 7-12. She is a member of the American Educational Research Association, Science Teachers' Associations of New York State and Phi Delta Kappa, the International Honor Society for Professional Educators. She has authored numerous publications.

Dr. Paula Szulc Dominguez, Dr. Craig Nicholls, and Barbara Storandt may be reached at 1201 East Fayette Street, Syracuse, New York, 13210, phone: 315-422-3512, or email: info@hezel.com. For more information, please visit us on the web at: www.hezel.com/resources.

Bibliography

Ball, D.L, Lubienski, S., & Mewborn, D. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematics knowledge. In V. Richardson (Ed.), *Handbook of research on teaching,* 4th edition, 433-456. New York: Macmillan.

Darling-Hammond, L., Holtzman, D., Gatlin, S.J. & Helig, J.V. (2005). *Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness.* Stanford University, April 2005. Retrieved online March 29, 2006 from http://www.schoolredesign.net/binaries/(teachercert.pdf.)

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Guskey, T. (1986). Staff development and the process of teacher change**. *Educational Researcher, 15*(5), 5-12.**

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2005). *The new instructional leadership: Creating data-driven instructional systems in schools*. WISC Working Paper Number 2005-9.
Retrieved online March 29, 2006 from
http://www.wcer.wisc.edu/publications/workingPapers/Working_Paper_No_2005_9.pdf

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Thomas, W. (2004). *Online professional development: Why SREB states should use it*. Retrieved online April 4, 2006 from http://www.sreb.org/programs/EdTech/pubs/PDF/04T05-OnlineProfDev.pdf

Weiss, I.R. (1994). *A profile of science and mathematics education in the United States, 1993.* Chapel Hill, NC: Horizon Research Inc.

Wenglinsky, H. (2002). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10* (12). Retrieved March 28, 2006 from http://epaa.asu.edu/epaa/v10n12/.