2010

Modeling Teacher Professional Development and Classroom Implementation of Instructional Strategies for Building Scientific Classroom Discourse Communities

Elizabeth B. Lewis
University of Nebraska-Lincoln, elewis3@unl.edu

Dale R. Baker
Arizona State University, DALE.BAKER@asu.edu

Nievita Bueno Watts
Arizona State University

Brandon A. Helding
Arizona State University

Michael Lang
National Center for Teacher Education, Maricopa Community College District, mike.lang@domail.maricopa.edu

MODELING TEACHER PROFESSIONAL DEVELOPMENT AND CLASSROOM IMPLEMENTATION OF INSTRUCTIONAL STRATEGIES FOR BUILDING SCIENTIFIC CLASSROOM DISCOURSE COMMUNITIES

Elizabeth B. Lewis, Henzlik Hall, Department of Teaching, Learning, and Teacher Education, College of Education and Human Sciences, University of Nebraska-Lincoln, Lincoln, NE, 68588

Dale R. Baker, Payne Hall, Mary Lou Fulton Institute and Graduate School of Education, Arizona State University, Tempe, AZ, 85287

Nievita Bueno Watts, Payne Hall, Mary Lou Fulton Institute and Graduate School of Education, Arizona State University, Tempe, AZ, 85287

Brandon A. Helding, Farmer Hall, Mary Lou Fulton Institute and Graduate School of Education, Arizona State University, Tempe, AZ, 85287

Michael Lang, National Center for Teacher Education, Maricopa Community Colleges District Office, Phoenix, AZ 85281

Abstract

Three-hundred-and-twenty-three classroom observations of secondary science and language arts teachers were made over two academic years while teachers engaged in professional development (PD) in how to construct scientific classroom discourse communities. These observations were used, along with teacher demographic information, to build a hierarchical linear model to explore statistical relationships. The length of time that teachers received PD was chosen as the exclusive predictor of teacher change while a schools' percentage of students who qualified for free and reduced lunch (a proxy for SES) was chosen as the exclusive predictor of intercepts. Over the course of two years, the teachers who had participated for longer periods of time used more of the PD, that is, they had higher rates of change than newly participating teachers. The model indicated, with statistical significance, that SES predicted teachers' baseline levels of behavior associated with the PD they were yet to receive. However, with respect to teachers' change over time, only the amount of PD that a teacher received or their treatment group membership predicted use with statistical significance. Ergo, while teachers' students' SES

was important in determining where teachers began, the treatment itself accounted for how teachers' instructional practices changed over time.

Introduction

Teacher professional development in science education is commonplace, but little is known about the degree to which science teachers apply professional development (PD) to their own contexts. For example, a recent report that sampled PD initiatives for math and science teachers nationally from 2004-2007 reported on effective PD activities that changed teachers' classroom instruction practices were over 50 hours in length (CCSSO, 2008). The CCSSO (2008) report estimates that only a third of the sampled evaluation studies reported such measurable effects of teacher PD.

Teachers are vital actors in bridging the academic culture and language of science and students' everyday popular culture and personal identities that are influenced by gender, ethnicity, and socioeconomic status. Engaging in professional development and learning encompasses aspects of individual cognition, social interaction, and the learning environment itself. These variables are in a constant state of flux, which complicates studying the ways teachers: (a) learn from professional development programs, (b) reflect on their teaching practices, and (c) implement professional development in their classrooms. As a result, few studies have considered the interaction among teachers' professional development, their classroom practice, and student performance (Hewson, 2007). Ultimately, we need to better understand how information is used by teachers in situ: between professional development and teachers, and between teachers and students. This will assist professional development providers and support future science education reform efforts.

In this study, we expanded upon a prior analysis (Lewis, Baker, Lang, & Helding, 2009) of one academic years' worth of classroom observations by following participating teachers for a second year to discover what relationships exist among teacher- and school-related variables and enacted PD. These new data were generated by observing a subsample of the previous 2007-2008 academic participants and additional teaching colleagues recruited into six, school-based teams.

Relevant Literature

Teacher learning and professional development. As Borko (2004) reported, "we have evidence that professional development can lead to improvements in instructional practices and student learning" (p. 3). In this study, specifically, our main objective was to understand teachers' application of the PD model as they designed and implemented curriculum and instruction to construct scientific classroom discourse communities. In this study's particular professional development program design, the instructional strategies were carefully selected from the relevant research literature to optimize learning science concepts, but what teachers decided to use with their students ultimately determined what, if any, benefits students gained as a result of teachers' participation in the PD.

Fidelity of implementation. O'Donnell (2008) discussed extant research on how teachers use professional development in classrooms, "there are too few studies to guide researchers on how fidelity of implementation to core curriculum interventions can be measured and related to outcomes, particularly within efficacy and effectiveness studies, where the requirements for fidelity measures differ" (p.33). While teacher professional development holds the promise of improving students' understanding of science, without teacher fidelity to reformed-based professional development models we cannot expect changes in student achievement. Without understanding what variables influence teachers' decisions to use, or not use, professional development as intended, the endeavor of professional development can only be a trial-and-error process.

The myriad reasons why or why not teachers choose to implement professional development, is in this work, a matter of understanding the effectiveness of professional development with coherent and theoretically sound statistical models. The challenge of determining the effectiveness of, and fidelity to, teacher professional development over time can be addressed with research methods such as longitudinal modeling (Shadish, Cook, & Campbell, 2002; Raudenbush & Bryk, 2002). Specifically, we used hierarchical linear modeling to identify what variables influence teachers' decisions to use, or not use, professional development as intended. The variables we considered were taken from previous work and are described below.

Socioeconomic status, class, and pedagogy. In her pivotal article on the hidden curriculum of work and social class, Anyon (1980) described how different social classes were subjected to different levels of pedagogy in her study of five schools of varying socioeconomic means. Teachers' instructional practices were affected by the socioeconomic status of their students. For example, in the working class community students' curriculum was dominated by activities in which they were expected to follow step-by-step procedures, while children of the affluent professional class were provided opportunities to develop skills to be the culture producers of the future. The later is more in line with science education reform and the process of scientific inquiry in which students ask questions design their own procedures for investigation, generate data, and make meaning from their academic science experiences. In science education, inquiry-based science instruction is the gold standard for all students, not just those who are perceived to be college-bound.

Professional Development & Research Context

The National Science Foundation-funded *Communication in Science Inquiry Project* (CISIP) provides school-based teams of secondary science and English and/or ELL teachers with year-round PD. The goal of CISIP was to teach secondary teachers how to build scientific classroom discourse communities (SCDC) with their students. The CISIP model included: (a) scientific inquiry; (b) oral discourse; (c) written discourse; (d) academic language development; and (e) essential learning principles such as accessing students' prior knowledge, forming conceptual frameworks, and incorporating metacognition (NRC, 2000; NRC, 2005). The use of scientific inquiry as a teaching platform provides students with opportunities to engage with scientific questions, make observations, and make meaning from their own experiences. To paraphrase Gee (2005), students need to experience science in order to be able to create meaningful discourse and develop conceptual understandings.

Middle and high school teachers participated in one of two three-week CISIP summer institutes in 2007. During the 2007-2008 academic year, four day-long PD workshops were held to build on the material that had been presented in the summer institute. The teachers had an opportunity to attend a total of 96 hours' worth of PD. However, some teachers had also previously participated in the development phase of CISIP and had potentially up to an additional 200 hours (over 2 years) of PD experience. In our follow-up study, only high school science and language arts teachers from two school districts were observed in the second year. These teachers participated in a four-day summer institute in July 2008 and engaged in 6 additional PD workshop days throughout the 2008-2009 academic year. This resulted in an additional 60 hours of CISIP PD contact time.

Research Questions

The main research questions for this exploratory investigation were:

1) Does the amount of PD a teacher receives significantly predict teacher implementation of the CISIP PD model or their baseline levels of PD-associated behavior?
2) Does the length of time that teachers have taught significantly predict CISIP implementation or baseline levels of PD-associated behavior?
3) Does the level that teachers teach, middle school or high school, significantly predict teachers' PD implementation or their baseline levels of PD-associated behavior?
4) Does the socioeconomic status of the school in which the teacher teaches significantly predict teachers' PD implementation or their baseline levels of PD-associated behavior?

Methodology

Considering the current challenge of determining the effectiveness of, and fidelity to, teacher PD over time, we used hierarchical linear modeling (HLM) to explore possible relationships between PD, teachers' practice, and other possible predictors of teacher change over time. Consequently, we attempted to build both two- and three-level HLM's using a range of systemic and professional development-related factors.

Data Collection

The 2007-2008 academic year was a pilot evaluation, after a two-year program product development phase in the CISIP grant. One data collection method was regular classroom observations of teachers' instructional practices. A second was a teacher demographic survey that gathered educational and professional information about each teacher. Some of this information, such as length of time teaching and length of involvement with the PD, was used in the construction of the model. Recent state-generated results from required state testing, school district size, per pupil spending on classroom and total costs, and socioeconomic variables (e.g., percentage of students eligible for free and reduced lunch) were also tabulated and reformed as variables for the model building process (Arizona Department of Education, 2008).

Classroom observations. The most labor-intensive aspect of the study was to observe teachers in their classrooms throughout the year. One-hundred-and-sixty observations were made from October 2007 to May 2008 by the university research team either in pairs, during the fall of 2007, or individually, in the spring of 2008. These observations included 28 classroom visits to a comparison group of 13 secondary science teachers between February and April 2008 as part of a smaller study that pre- and post-tested students' knowledge of genetics and heredity. Over the course of the 2008-2009 academic year we made an additional 163 observations of ten of the same teachers, seven teachers who had previously participated in the PD, but not the research, and 16 of their newly-recruited teaching colleagues.

The CISIP classroom observation instrument, the "Discourse in Inquiry Science Classrooms" (DiISC), formerly named the CISIP Classroom Observation Instrument (COI), was developed and refined to be aligned with the PD model over a period of four years (Ozdemir, Lewis, & Baker, 2007). Initially the items were developed with reference to the research literature base regarding the role of writing, oral discourse, scientific inquiry (NRC, 1996), learning principles in science teaching and learning (NCR, 2000, 2005), and academic language development strategies. We have not yet psychometrically verified the DiISC as we would need more observations to do so. It was, nevertheless, our outcome variable across the models we considered in answering our research questions.

Results

Participants & Data

A group of 23 secondary teachers who participated in the CISIP year-round PD were observed during the 2007-2008 academic year. There were 15 science and 8 English, ELL, and library media teachers in the sample who had taught from 0 – 32 years as of the summer PD. The comparison group of 13 science teachers was also observed during one unit of instruction in the spring of 2008. Observations were conducted using the DiISC with 36-items, each with individual rubrics that employed a 0 to 3 point Likert-scale. During the 2008-2009 academic year, we observed 30 teachers (16 science and 14 language arts).

For the two-year longitudinal model we prepared two and three levels of data. The first level for both models included the total raw observation scores on the DiISC. The second level included: a code for participation in the PD or comparison group, the total length of time (in months) that the teacher had participated in CISIP, the grade level (middle or high school) each teacher taught, the number of students attending each teachers' school, the school's overall performance on state-mandated tests, the number of students in the district, the classroom and total per pupil spending costs, the percentage of students eligible for free or reduced lunch at each school (SES), average teacher pay in the district, and the number of years the teachers had taught. A third level was attempted by separating the school and district-level variables, but in the end we did not have the sample size to converge on an estimated solution. That is, we did not have enough individual schools in the study to support so complex a model.

Data Analysis

In our initial study of the 2007-2008 school year data, we built a model that described, with statistical significance, the teachers' change in using the PD over time. The model equations that resulted were:

Level 1: PD Use $= \Pi_0 + \Pi_1(\text{Time}) + e$

Level 2:
$$\Pi_0 = \beta_{00} + r_0$$
$$\Pi_1 = \beta_{10} + \beta_{11}(\text{PD Length}) + r_1$$

Because the model was ultimately based on the raw metric, and lacked a fully comparable control group, the conclusions we made were not generalizable to other PD programs or other groups of participants. We determined from this that further investigation was necessary, especially as there was significant variance in both the intercept and the slope to be further modeled (when evaluated using residual variances and an alpha of 0.05, see previous work for specific variance components). In our follow-up analysis of two years' worth of observation data we designed two two-level HLM's using the total raw DiISC measures. The two models used the following equations:

Model A

Level 1: PD Use $= \Pi_0 + \Pi_1 *(\text{Time}) + e$

Level 2:
$$\Pi_0 = \beta_{00} + \beta_{01} *(\text{Poverty}) + r_0$$
$$\Pi_1 = \beta_{10} + \beta_{11} *(\text{Condition}) + r_1$$

Model B

Level 1: PD Use $= \Pi_0 + \Pi_1 *(\text{Time}) + e$

Level 2:
$$\Pi_0 = \beta_{00} + \beta_{01} *(\text{Poverty}) + r_0$$
$$\Pi_1 = \beta_{10} + \beta_{11} *(\text{Total PD}) + r_1$$

Both models fit similarly well, or less bad, after all data were considered and interactions terms eliminated as possible predictors. Both models had significant predictors of intercept and slope. The actual predictors of slope, however, differed; that is, they were different conceptualizations of treatment. In Model A, treatment was a simple 1 or 0, grouping value. In Model B, that group membership was broken into the actual amount of PD that any one teacher received. The overall model calculations for Model A are summarized in Table 1. Figure 1 is a simplified graph of Model A according to multiple levels of SES and treatment condition; by SES at 14%, 33%, 47%, 57%, 70%, and 95% and by participation in the PD (dashed lines) or not (solid lines). In Figure 2, slopes were constrained in order to demonstrate how the starting points of teachers varied across levels of SES. In Figure 3, the teachers' starting points/intercepts were constrained in order to demonstrate how the slopes varied across levels of treatment.

Table 1

Overall model.

	Effect (variable)	B	*Se*	*t* Ratio	d*f*	*p*-value
Intercept, Π_0						
	Intercept, β_{00}	38.22	4.70	8.13	58	< 0.01
	Poverty	-19.48	7.96	-2.45	58	0.018
Slope, Π_1						
	Intercept, β_{10}	-0.012635	0.011984	-1.05	58	0.297
	Condition, β_{11}	0.023016	0.009537	2.41	58	0.019

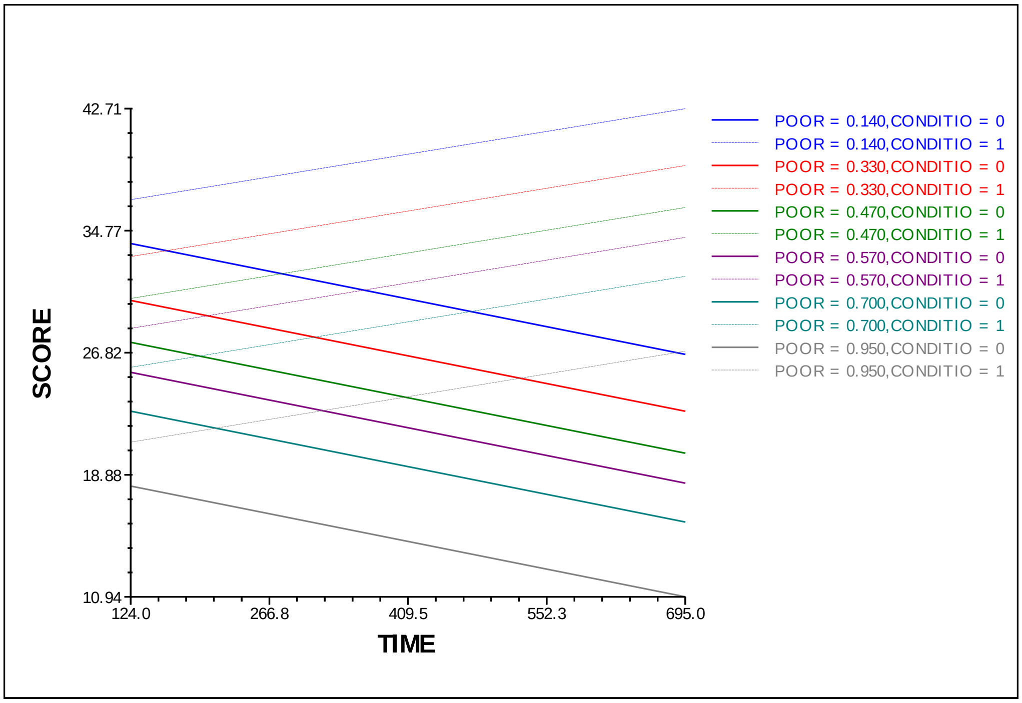

Figure 1. Model A: Multiple levels of SES and professional development. The solid negative-sloping regression lines represent teachers in the non-PD comparison group with intercepts based on the SES of their student populations. The dashed positive-sloping regression lines represent the CISIP teachers engaged in professional development grouped by students' SES.

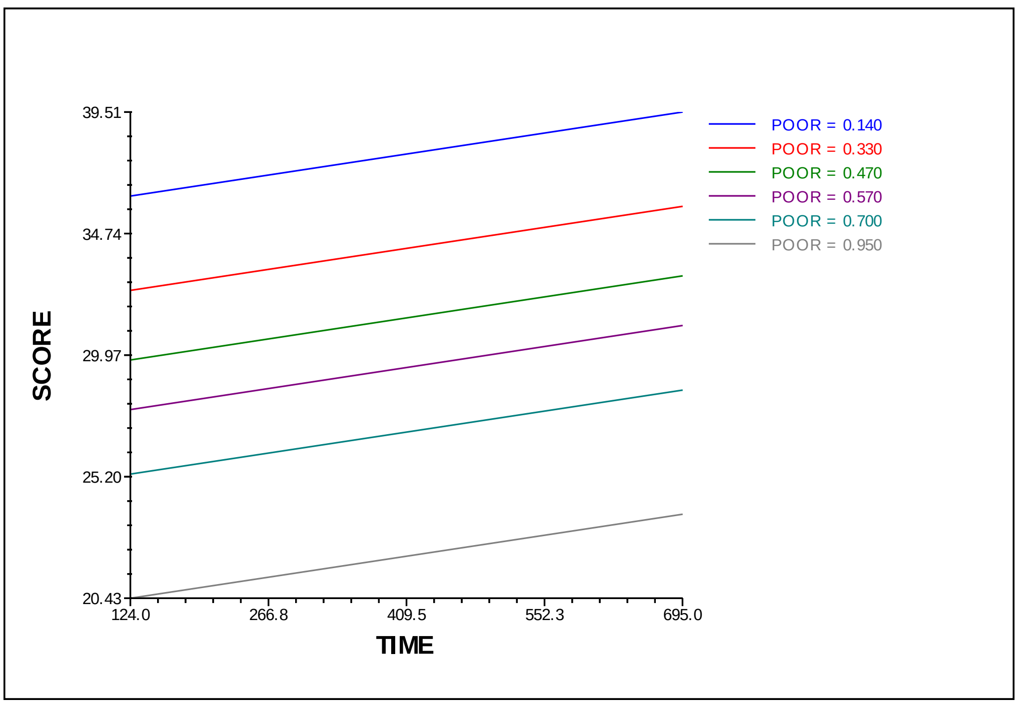

Figure 2. Model A: holding slopes constant, with intercepts varying by SES. The percentage of students qualifying for free and reduced lunch decreases from 95% on the lowest regression line to 14% at the highest. In the model all teachers increase in their use of CISIP instructional strategies, but begin at a range of scores based on the SES of their students.

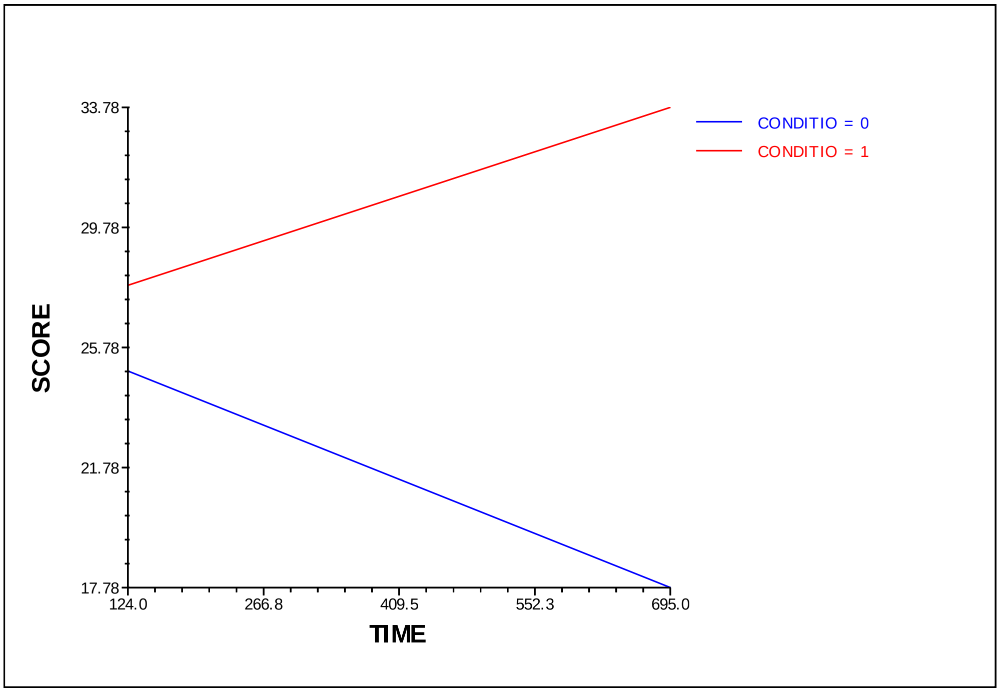

Figure 3. Model A: Professional development vs. no professional development. Over time teachers engaged in the professional development increased in their use of the CISIP instructional strategies while the group of comparison teachers without CISIP professional development decreased in their use of CISIP instructional strategies.

Discussion

The length of time that the teachers received PD was chosen as the exclusive predictor of teacher change while a schools' percentage of students who qualified for free and reduced lunch was chosen as the exclusive predictor of the intercept/starting point. Over the course of two years, the teachers who had participated for longer periods of time used more of the PD and had higher rates of change than newly participating teachers. The model indicated, with statistical significance, that SES predicted teachers' baseline levels of PD-associated behavior. However, with respect to teachers change over time, only the amount of PD that a teacher received or their treatment group membership predicted use with statistical significance. Ergo, while teachers' students' SES was important in determining where teachers began, the treatment itself accounted for how teachers changed over time.

Also, consider the multileveled regression lines in Figure 3. On a long enough time line, the comparison group teachers' PD-associated behaviors would become negative and the treatment group teachers' PD-associated behaviors would approach infinity. But the CISIP measure has no meaningful negative or very large values. The linear nature of the relationship, outside the range of our data, is therefore de facto absurd. This indicated that, although our models fit tolerably well, they would not apply outside the range of our data. More sensitive data collection schedules, and associated model building processes, would likely require some other linear form (our models in this case are only curvilinear) (i.e., logistic hierarchical linear modeling). Additionally, the CISIP measure itself, in future work, will need psychometric verification so that the change over time it notes can be extended into work that does not use it and only it as an outcome measure.

Lastly, while we found evidence that the PD was associated with teacher change, these claims are tentative and subject to further verification with more rigorous research designs and analyses. Because SES significantly predicted teachers' initial use of CISIP PD instructional strategies further research, both by the author(s) and others, will and should include more teachers and/or schools to fully determine how SES interacts with the complexity of schooling and instructional practices.

Conclusion

As a result of the HLM and the inspection of the DiISC data, our conclusions are limited to the scope of the DiISC, with all of its strengths and limitations. Long-term CISIP PD appears to predict greater implementation of teaching strategies that foster the development of scientific classroom discourse communities. Those who design and or study professional development should bear in mind the importance of observing teachers frequently over long periods of time in order to accurately model teacher change over time as well as document the sometimes delayed effects of professional development. Additionally, if the goal is to find results that are generalizable outside the scope of the study itself, the first task in future investigations will be to revise the DiISC and account for discrepancies in the sampling procedures and ultimately to eliminate plausible, alternative hypotheses in search of causal links.

Acknowledgements

This research was supported by the National Science Foundation, grant # 03353469.

References

Anyon, J. (1981). Social class and school knowledge. *Curriculum Inquiry, 11(1)*, 3-42.

Arizona Department of Education (2008). Arizona Department of Education School Report Cards 2006-2007. Retrieved July 12, 2008. Web site: http://www.ade.az.gov/azlearns

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33(8),* 3-15.

Council for Chief State School Officers (2008). *Does Teacher Professional Development Have Effects on Teaching and Learning?* Washington D.C.

Gee, J.P. (2005). Language in the Science Classroom: Academic Social Languages as the Heart of School-Based Literacy. In R.K. Yerrick & W.-M. Roth (Eds.), *Establishing Scientific Classroom Discourse Communities: Multiple Voices of Teaching and Learning Research* (pp. 19-44). New Jersey: Lawrence Erlbaum Associates, Publishers.

Hewson, P. W. (2007). Teacher professional development in science. In, S. K., Abell, & N. G., Lederman (Eds.). *Handbook of Research on Science Education* (pp. 1177- 1203). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Lewis, E.B., Baker, D.R., Lang, M., Helding, B.A. (2009). Using HLM to Analyze On-Going Teacher Professional Development and Implementation of Scientific Classroom Discourse Community Strategies. Paper presented at the annual meeting of the National Association for Research in Science Teaching: Garden Grove, CA.

National Research Council (2005). *How Students Learn: History, Mathematics and Science in the Classroom, A Targeted Report for Teachers*. M. Donovan and J. Branford (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academy Press.

National Research Council (2000). *How People Learn.* J.D. Bransford, A.L. Brown, and R.R. Cocking (Eds.). Washington, DC: The National Academy Press.

National Research Council (1996). *National Science Education Standards*. Washington, DC: The National Academy Press.

O'Donnell, C. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78(1)*, 33-84.

Ozdemir, G., Lewis, E.B., Baker, D.R. (2007). Development and Validity of the CISIP Classroom Observation Instrument (COI). Paper presented at the annual meeting of the National Association for Research in Science Teaching: New Orleans, LA.

Raudenbush, S., & Bryk, T. (2002). *Hierarchical linear models: applications and data analysis methods.* Thousand Oaks, CA: Sage Publications.

Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.