Impact of Professional Development in an Interdisciplinary Science and Engineering Partnership on Teacher's Knowledge and Practice and Student Understanding of Science

Yang Yang*
Xiufeng Liu
Michelle R. Eades-Baird

*Graduate School of Education
University at Buffalo, SUNY
e-mail: yang37@buffalo.edu

Paper presented at the annual meeting of the NARST – A World-wide Organization for Promoting Science Teaching and Learning through Research, April 17, 2016, Baltimore, MD

Acknowledgement: This paper is based upon work supported by the National Science Foundation under Grant No. DUE-1102998. Any opinions, findings, and conclusions or recommendations expressed in the materials are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Impact of Professional Development in an Interdisciplinary Science and Engineering Partnership on Teacher's Knowledge and Practice and Student Understanding of Science

Introduction

Calls for improving students' science achievement through standards-based science education reforms are unprecedented in the past few decades, from *Science for All Americans* (Rutherford & Ahlgren, 1990) to the present frame work of *Next Generation Science Standards* (NGSS) (NRC, 2012). The focus of *Science for All Americans* was on blurring the boundaries between science disciplines by stressing connections between traditional science subjects and emphasizing more conceptual understanding and thinking skills rather than the memorization of facts and procedures (Bybee, 1995). The NGSS took another major step forward through its expectations of student performance in terms of three interconnected dimensions: science and engineering practices, crosscutting concepts, and disciplinary core ideas (NRC, 2013). The NGSS is deeply rooted in interdisciplinary science inquiry, learning sciences, and science education research, and has been well endorsed by both researchers and practitioners. The concept of interdisciplinary science inquiry in the new standards is consistent with previous definitions of interdisciplinarity that it is a blending of sciences where connections are made between the subjects, but they remain identifiable (Czerniak, 2007).

Teacher's knowledge and practice

Despite of good intentions of new curriculum standards, science education reform faces great challenges in terms of implementation. At the classroom level, implementation of the standards relies mostly on teacher's knowledge and practices. According to Gess-Newsome and Lederman (2001), teacher's knowledge consisted of five domains, which include 1) knowledge about the general educational context, 2) knowledge about the specific educational context, 3) general pedagogical knowledge, 4) subject matter knowledge, and 5) pedagogical content knowledge. The last two types of knowledge have drawn most attention in in-service teacher professional development and research. Subject knowledge refers to teacher's understanding of subject matter taught while pedagogical content knowledge is the requirement to make the subject matter accessible to the students (Shulman, 1986). The interdisciplinary science understanding of teachers, for example, the integration with Nature of science and new technology, would improve a teacher's ability in creating sociologically authentic science experience (Cunningham, 1998). Both content knowledge and pedagogical content knowledge are key components of teacher competences that affect student progress (Kleickmann et al., 2013).

Teacher knowledge and beliefs play a significant role in their decision-making about curriculum and instructional tasks (Nespor, 1987). A large number of studies have been done on the relationship between teachers' knowledge/beliefs and their classroom practices. For example, that a teacher's understanding of the nature of science is reflected in classroom practice is not automatic and is extremely complex in nature (Abd-El-Khalick et al., 1998), and in another study of teachers' knowledge/beliefs in applying science inquiry, the authors found that inquiry-based teaching will be difficult to implement in the current school culture (Wallace & Kang, 2004). Because teacher knowledge and classroom practice are major mechanisms to improve student science learning (Wilson, 2013), the development of science teacher knowledge and practices is essential. In-service teacher professional development (PD) is a common way in scaffolding such development.

Teacher professional development

According to van Driel (van Driel et al., 2012), PD refers to the procedures and activities designed to consolidate teacher professional knowledge, skills and attitudes and to further improve student achievement. A growing number of studies illustrate a positive effect of PD on students learning outcomes (Blank et al., 2007; Cohen & Hill, 2000; Desimone, 2009; Franke et al., 2001; Roth et al., 2011; Saxe & Gearhart, 2001; Scher & O'Reilly, 2009). However, a few researchers still find the possibility of negative effect of PD on students in the past two decades (Aitken & Sinnema, 2008; Nuthall & Alton-Lee, 1995; Seixas, 2001), which reflects the issue of quality and implementation of PD. Therefore, high quality PD, when aligned with objectives of reform, could be a powerful method to promote academic innovation. It is widely accepted in research that effective PD in science education consists of six components including content focus, active learning, collaborative learning, duration, coherence, and school organizational conditions (Borko, 2004; Heller et al., 2012; van Driel et al., 2012; Wilson, 2013).

Nevertheless, the quality or the relevance of PD is not typically regulated in mandatory programs for license renewal each year and effective PD design supported by empirical evidence is needed (Grossman & Hirsch, 2009). Furthermore, the mechanism of impact of PD on students is less studied. Heller and colleagues (Heller et al., 2012) reported how three different interventions in science PD affected teachers and students' learning outcomes. Diamond et al (Diamond et al., 2014) studied the influence of PD on elementary teachers' science content knowledge and their students' science achievement. Their work shed the light on relationship between a variety of features in PD and learning outcomes. But there are a large number of questions remaining unanswered in this area.

It is reasonable to say that a well-designed PD program could directly improve teachers' profession in a myriad of ways, which could include pedagogical content knowledge, content knowledge, attitudes and beliefs towards science education, classroom instruction, understanding of nature of science, and so on and so forth (Gess-Newsome & Lederman, 2001). Although there are a number of empirical studies on teacher PD, few provided evidence on this relationship. Furthermore, research has limited to one grade level and one subject content area; few studies addressed multiple grades with a focus on interdisciplinary science inquiry. Finally, research using reliable and validate measurement of student achievement in understanding interdisciplinary science was limited. Thus, more work is needed to extend the scope of such studies in aspects of grade level and interdisciplinary field. In addition, a more reliable and validate measurement is required to improve the quality of research.

Students' learning outcome in science

Large-scale standardized assessments are the most widely used for evaluating student learning outcomes in science, which is necessary for educational policy making (Britton & Schneider, 2007). Small-scale assessments, such as school-wide or classroom tests, connect with everyday teaching and learning more strongly. Although some alternative assessment methods, such as interview and performance assessment (Klassen, 2006; Mintzes et al., 2005), can have high reliability and validity and could measure student science understanding, multiple-choice questions are still the most common choice in any kind of exams (Roediger III & Marsh, 2005).

Accordingly, majority of research, when studying student learning outcomes in science, adopt test scores as the indicator.

According to Metz (Metz, 1998), a student's change in science knowledge was influenced by their prior knowledge, naïve philosophy of science, characteristic of input information, and processing strategies. Therefore, identification of students' prior misconceptions, beliefs and attitude towards science, and learning strategies are essential elements in studying studentlearning outcome in science. Understanding of Nature of Science (NOS), as a part of student naïve philosophy of science, affects their learning outcome in various ways. For example, Songer (Songer & Linn, 1991) found that the belief that science is an enterprise of accumulating facts rather than constructing theories to account for data, could inhibit students' learning of new theories. Confidence is also a potentially powerful influence on student achievement (Lau & Roeser, 2002).

Furthermore, teacher expectations have been found to be one contributor to achievement gaps of students of different races (Weinstein et al., 2004). In addition, integrated inquiry and discourse have been studied in science education (Bybee et al., 2008), and the mean effect size is 0.5 in a meta-analysis of empirical studies from 1996 to 2006 (Furtak et al., 2012). This study involved the aforementioned influences on students' science achievement, which include student self-efficacy, understanding of Nature of Science, teacher attitude/expectation on student work, and teacher support in discourse/inquiry.

Interdisciplinary Science and Engineering Partnership (ISEP)

This research study took place within the context of the Interdisciplinary Science and Engineering Partnership (ISEP), a National Science Foundation (NSF) funded 5-year professional development program. The aim of this program was to improve the quality of science teaching and learning through promotinginteractions between science, technology, engineering, and mathematics (STEM) in twelve low-performing public schools within a large urban school district in the Northeastern United States. The program consisted of a myriad of activities that include school based wrap around support, teacher professional development, and assistance of STEM graduate students in classrooms.

The ISEP adopted the PD model (Desimone, 2009) by incorporatingsix aspects of high-quality professional development. The major activities of PD were summer experiences of teachers and monthly professional learning community (PLC) sessions focusing on pedagogy. The former contained three sub-categories: interdisciplinary research, science curriculum study, and college courses in physics and engineering. The goal of summer research experience for teachers was to improve science teachers' knowledge and skills in interdisciplinary science inquiry through engagement in science research and engineering design with university STEM faculty. Curriculum study, facilitated by university STEM faculty, helped teachers develop resources to enhance their courses. The summer courses on physics and engineering were offered by a local state college. The monthly workshop offered by the educational research team focused on interdisciplinary science inquiry pedagogy. Finally at the end of each academic yearteachers were required to make a poster presentation on their implementation of interdisciplinary inquiry in their schools and classrooms.

The school-university partnership is beliefed to be an effective way to carry out teacher in-service PD (Cardullo & Forsythe, 2013; Pepper et al., 2012; Shroyer & Yahnke, 2012). The partnership could bring interdependence and mutual benefits to both schools and universities. Although few articles show issues in the partnership, such as identity conflict of the trainees (Trent, 2012), the majority studies support a very positive outcome out of the partnership (Cheng & So, 2012; Douglas, 2011; la Velle et al., 2013; Whitehead & Fitzgerald, 2006). During the partnership, schoolteachers had access to the rich resource in the universities, subject courses, research opportunities, communication with experts, just name a few. It provided a great opportunity to build a development and research cycle for the teachers, which is considered as the key to effective PD (Alton-Lee, 2011).

The purpose of the study is two-fold. First, the authors will explore possible relationships among PD intervention, teachers' PCK and classroom practice, and students' learning outcome within the project. Second, the study will provide empirical evidences of effectiveness of PD intervention on teachers' PCK/practice and students' understanding of science. The research questions guiding this study are:

(1) What is the relationship, if any, between PD intervention and results of teachers' PCK assessment?
(2) What is the relationship, if any, between PD intervention and student report of teacher classroom practice?
(3) What is the relationship, if any, between PD intervention and students' science achievement in terms of their ability in understanding of interdisciplinary science?
(4) What is the mediation effect, if any, between PD intervention and students' science achievement through teachers' PCK/classroom practice and students' attitude/beliefs in science?

Literature review

The PD has direct influence on teachers' knowledge, beliefs, and attitude towards teaching their own subjects, which would further affect students' learning outcomes. Diamond (Diamond et al., 2014) proposed a model for the impact of PD on teacher science content knowledge and student achievement. Similarly, a model of the effect of teachers' PCK gained from PD on students' science achievement is shown below:

The measurement of teachers' CK and PCK is various, including teacher self-reported confidence in teaching a certain science topic, content test, class observation, and so forth (Diamond et al., 2014; Heller et al., 2012; Roth et al., 2011). Furthermore, some indirect indicators are also proved to be useful, such as amount of relevant courses taken in college and teaching qualification (Ball, et al., 2001; Nowicki et al., 2013). The indirect methods are always adopted since the cost of test and observation are too high. To measure students' science achievement, survey and standardized test scores are most commonly used.

Stolk (Stolk et al., 2011) provided evidence that designing new context-base unit in chemistry according to the curriculum only partially empowers teachers. The number of relevant college courses is found to correlate with teachers' self-reported CK (Diamond et al., 2013; Shallcross et al., 2002). Furthermore, research-based programs could provide an opportunityfor the teachers to join a real scientific investigation, which was aligned with the purpose of inquiry science. Teachers then would have great opportunities to join the hands-on activities and communicate with the experts. Seraphin (Seraphin et al., 2013) studied influence of online PD on teaching science as inquiry framework, and found an in-depth use of the framework phases would have more effect on the teachers' instructional method towards their students. Tal and Argaman (Tal & Argaman, 2005) studied teachers participating in an in-service PD that was providing both SCK and PCK in science inquiry activities. They found the inexperienced teachers lacked confidence and ability in managing their students in the inquiry based activities compared with the experienced teachers.

Garet (Garet et al., 2001) and colleagues examined the relationship between features of PD that have been identified in the literature and self-reported change in teacher's knowledge & skills and classroom teaching practice. They found sustained and intensive PD was more likely to have an impact than short ones. In addition, subject matter, hands-on work (active learning), and integration to daily life of school were more likely to produce enhanced knowledge & skills. Moreover, type of PD had indirect effect through duration and core features. Although the importance of collective participation was shown, which involved teachers from same school, subject, or grade working together, the relationship was not significant. Supovitz (Supovitz & Turner, 2000) and colleagues conducted a similar study with focuses on length of PD, teacher beliefs and background, and school demographics. Their results showed that 80 hours PD could significantly change classroom practice while it required 160 hours to change classroom culture. Furthermore, teacher content knowledge preparedness was a powerful predictor of teacher classroom practice and culture. At the school level, poverty in terms of percentage students with free or reduced lunch was the only significant parameter. It meant teachers from poor schools appeared to use more traditional teaching practice. In another study of relationships between the extent of 8th grade teacher participation in a PD program and their attitudes, perceptions of content & pedagogical preparedness, and classroom practices, the authors found the PD hours and principal support were positively related to teacher attitudes, preparedness, investigative practice, time devoted to instruction, application of instructional materials, and investigative practice (Banilower et al., 2007). They also found the application of instructional materials and time devoted increased dramatically at 100 hours of PD. The mediation effects happened through attitude, preparedness, and application of instructional materials between PD and classroom practice.

In Diamond et al.'s study aforementioned (Diamond et al., 2014), they examined the effect of a PD intervention on teachers' SCK/PCK and teachers' knowledge on student learning outcome in terms of high stake science test scores. The results illustrated higher achievement in the treatment group of teachers compared with control groups in both SCK/PCK tests. In the further multi-level analyses, teachers' years of experiences had a small but positive effect on student science achievement. Teacher participation in PD that directly targeted their SCK at the grade level may have a direct positive effect on student achievement, though the variances explained by the two predictors were small (3% and 6% respectively). Lee (Lee et al., 2008) studied the effects of PD on student science achievement gaps and growth, especially for ESL students. The PD emphasized in this research included English, home language and culture besides of science content knowledge. The students demonstrated significant gains at the end of each school year from grade 3 to 5. The achievement gap narrowed with 4th graders sometime and remained consistent with 3rd and 5th graders. Achievement gains (growth) and effect size tended to be lower on

NAEP/TIMSS tests than those on the project-developed units because the latter was closely aligned with the PD.

Method

Design Overview and sample

The data were collected from five sources as shown in the table below, teacher PD record and survey from summer 2012 to 2014, teacher SCK test in summer 2013 and 2014, student survey in three successive semesters (Spring 2013, Fall 2013, and Spring 2014), and student SCK test in Fall 2013 and Spring 2014. The data from summer 2014 to 2016 are still under processing from the external evaluator, and will be used in future study.

Table 1 Timetable of research design

	Summer 2012	Fall 2012	Spring 2013	Summer 2013	Fall 2013	Spring 2014	Summer 2014
Teacher PD record	✔	✔	✔	✔	✔	✔	✔
Teacher SCK test				✔			✔
Student survey			✔		✔	✔	
Student test					✔	✔	

This study used data of teacher PD from Fall 2012 to summer 2014, teacher SCK tests, student survey and tests in Fall 2013 and Spring 2014 to conduct the research. Because the research contained two separated analyses on teachers and students/teachers respectively, the data applied was different but overlapped. In the analyses of relationship between PD and teacher SCK test results, the data of all 93 teachers from 12 schools was adopted. While in the analyses of influence of PD on student science understanding, only elementary/middle school students and their teachers were focused. Therefore, the data in the later consisted of 696 students from grade 4 to 8 in 23 classrooms nested in 5 schools. The sample descriptions of both teachers and students are shown in the results.

Measures

Teacher SCK tests: the participating teachers were required to finish a SCK test related to their teaching areas. The test consisted of 6 different subject domains. The chemistry test assessed teachers' knowledge of teaching properties and changes in matter and it was developed by the Assessing the Impact of the MSPs: K-8 Science (AIM) project at Horizon Research, Inc.. The biology test measured teachers' understanding of the flow of matter and energy for teaching. The earth science assessment measured earth science teachers' understanding of plate tectonics for teaching. The physics/engineering instrument measured teachers' understanding of force and motion. The biology, earth science, and physics PCK assessment tools were developed by the Assessing Teacher Learning About Science Teaching (ATLAST) project at Horizon Research, Inc. The average reliability of the above tests in the study was 0.84. The elementary/middle school science assessments consisted of 8 items from POSTT "Thinking About Science Teaching" that related to teaching science to grades 1 through 8 (Cobern et al., 2014).

Student understanding of interdisciplinary science: the instrument contained 17 multiple-choice items from three sources, namely, *Science Attitudes, Skills, & Knowledge Survey (SASKS): Form 1-3, Discovery Inquiry Test (DIT), and Ohio Achievement Tests: Grade 5 Science Student Test Booklet.* The raw scores were calibrated through grade 4 to grade 8 in Fall 2013 and Spring 2014 by using Rasch measurement. The achievement of the students used in the study is Rasch ability of individuals. Validation and detailed analyses of the teacher and student tests could be found in another study of the authors.

Teacher PD record: teacher PD program included two major parts as aforementioned: summer placement and professional learning community (PLC). Summer placement consisted of research opportunities with STEM faculty in the university and a three-credit course in a local college. Teachers could choose a 4- or 6-week research project from 9 areas or the summer course in summer 2013, and in summer 2014 the curriculum study option with STEM faculty was added.   The features of the summer placement contained types and length of the activities. The PD hours were recorded as 6 hours per day for any types of placement to keep consistent, if the teacher took place on that day. For example, if a teacher joined a 4-week research project with full attendance, the PD hours would be 120 for the teacher. The attendance of PLC session provided the actual total hours of teachers in the monthly PLC sessions during the whole year.

Student survey: the survey consisted of 4 sets of questions, namely, my opinion about science (12 items), what teachers do in classroom (12 items), what students do in classroom (12 items), and parental support at home (7 items). The exploratory factor analysis showed three dimensions in the first set of questions, and two dimensions in the other questions. After discussion of the content of each item between two researchers and a confirmatory analysis, the results of the survey that reflected nine aspects are shown in Table 2. According to the purpose of the study, the latent traits generated from question set one and two were used in the analyses.

Table 2 Content of student survey

Subscale	Number of items	Factor	Items	Reliability
Q1. My opinion about science	12	Self-efficacy in science	Q8a, Q8b, Q8c, Q8d	0.76
		Use of science	Q8e, Q8f	0.38
		Understanding of the Nature of Science	Q8h, Q8i, Q8k, Q8l	0.61
Q2. What teachers do in classrooms	12	Teacher support in questioning and discussion	Q9a, Q9d, Q9f, Q9g, Q9h	0.78
		Teacher’s attitude & expectation on student’s work	Q9e, Q9i, Q9j, Q9l	0.63
Q3. What students do in classrooms	12	Learn science reasoning	Q10a, Q10d, Q10g, Q10h, Q10i, Q10j, Q10k	0.84
		Learn to corporation	Q10b, Q10c, Q10l	0.70

Q4. Parental support at home	7	Parental assistance	Q11a, Q11c, Q11d	0.82
		Parental expectation	Q11e, Q11f	0.68

Missing data was deleted under two criteria. For student test of understanding of science, the bar is set at 20% response missing. In other words, students who left more than 14 items blank would be deleted from the analyses, and in other situations, the blanks were considered as incorrect answers. In this step, 17% of the responses were deleted. For student survey questions, the case was deleted when the students answered less than 80% of the questions, only 7% of the sample was excluded and the rest missing responses were replaced by series mean. Finally, 509 students in 23 classrooms were used in the multi-level analyses.

Variables and data analyses procedures

To address the first research question, one-way ANOVA and a two-level hierarchical linear model (HLM) were applied because the teachers were nested in 12 schools, which might not share the same demographics (Raudenbush, 2004). Teacher test scores in summer 2013 and 2014 were contrasted. Then a model of teacher test scores with a series of parameters was built on two time points separately. The equation is shown as follow:

$$Y_{tij} = \beta_{t0j} + \beta_{t1j}X_{t1ij} + \beta_{t2j}X_{t2ij} + \beta_{t3j}X_{t3ij} + r_{tij} + u_{t0j}$$

Where Y_{tij} is SCK percentage scores of teachers; X_{t1} is a group of teacher demographical variable, including gender, experience, and highest degree; X_{t2} is PD feature, which includes summer placement type, attendance of PLC session, and total PD hours; and X_{t3} is a group of school level demographics that contain turnover rate, percentage of minority, percentage of poverty, and percentage of limited English proficiency.

To answer questions three, another two-level HLM model was built as followed:

$$Y_{sij} = \beta_{s0j} + \beta_{s1j}X_{s1ij} + \beta_{s2j}X_{s2ij} + \beta_{s3j}X_{s3ij} + r_{sij} + u_{s0j}$$

Where Y_{sij} is the student's Rasch ability; X_{s1} is a group of student demographics, namely, gender, race, and grade, which are treated as control variables; X_{s2} is a group of student level parameters that contain self-efficacy, understanding of NOS, student reported teacher expectation/attitude on their work and teacher support in science inquiry/discourse; X_{s3} is a group of teacher level predictors, which include teacher PD features, SCK results, and school demographics. School background was merged into the teacher level because of the small sample size of teachers in each school, and a pilot analysis also showed none-significant intra-correlation among school level.

Research questions two and four were tested by inter-correlations and regressions by using teacher classroom practice/student self-efficacy/NOS as outcome and teacher-level PD features and controls as parameters. The treatment of putting teacher-level predictors into student level because of the sample size illustrated another limitation of the study in this step.

Results

Descriptive statistics and SCK studies of teachers

Overall, 93 teachers in 12 public schools participated in the PD program from 2012 to 2014. Twenty-six teachers reported their gender and race, among which the male and female teachers were about the same and the majority was white with only one black teacher. In summer 2012, 44 teachers took part in disciplinary science research project at the university and 15 teachers joined science course in the local college. While in summer 2013, 50 teachers participated in the research projects, 10 joined the science course, and 14 worked in curriculum study. The descriptive statistics of the teachers' PLC attendance and SCK results are shown in Table 3. There were 44 teachers who took both summer 13 and 14 tests and a paired t-test was conducted on these teachers. There is no significant difference in both PLC attendance and PCK test results between two academic years from both ANOVA and paired t-test.

Table 3 Descriptive statistics of PLC and SCK tests

		PLC (hours)			SCK test (%)	
	N	Mean (S.D)	Range	N	Mean (S.D.)	Range
2012-2013	31	7.1 (4.0)	3-18			
Summer 13				66	64.6 (18.3)	24-100
2013-2014	30	5.8 (3.8)	1-13			
Summer 14				71	69.4 (18.4)	17-100

Because too fewer teachers reported their background, the predictors of teacher demographics were eliminated from the analyses, which illustrated another limitation of the study. The inter-correlations showed no significant relationships between teachers' SCK scores in summer 2013 and the other PD parameters, which include total PD hours in 2012 to 2013 academic year, PLC session participation hours in 2012 to 2013 academic year, and summer 2013 activities. The correlations in the other academic year showed similar results. Furthermore, the two-level HLM also showed non-significant results of all school level parameters. However, PLC attendance time in 2013 to 2014 academic year was positively related to teacher's PCK test scores in summer 2014, and the regression ($B = 1.283$, $p < 0.05$) coefficient from a simple regression illustrated that teachers with one hour more in PLC session than average scored 1.283 percent higher in their PCK test.

Descriptive statistics student's level parameters

The data of students and teachers in 2013-2014 academic year were used for HLM analyses. The sample contained 509 students in 23 classrooms in 5 elementary/middle schools. The descriptive statistics are shown in Table 4. Gender distributed evenly while for grade and race, there was a too small sample size in certain categories. For example, Asian students only took 5% of all sample students. The correlation between student average self-efficacy in science and understanding of NOS was significant ($r = 0.302$, $p < 0.01$). The correlation between student reported teacher support on science inquiry/discourse and teacher attitude/expectation on their work was also significant ($r = 0.460$, $p < 0.01$, see table 6).

Table 4 Descriptive statistics of students/teachers

Variable Name	N	Mean	SD	Minimum	Maximum
Student level					
Rasch ability	509	-0.20	0.82	-2.33	3.31
Self-efficacy	509	3.64	0.73	1.00	5.00
Nature of science	509	3.69	0.67	1.00	5.00
Teacher attitude/expectation	509	4.04	0.61	1.00	5.00

Teacher support	509	3.43	0.80	1.00	5.00
Gender	509	0.52	0.50	0.00	1.00
Grade 5	509	0.27	0.44	0.00	1.00
Grade 6	509	0.17	0.37	0.00	1.00
Grade 7	509	0.17	0.38	0.00	1.00
Grade 8	509	0.28	0.45	0.00	1.00
American Indian	509	0.03	0.17	0.00	1.00
Asian	509	0.05	0.21	0.00	1.00
Black	509	0.31	0.46	0.00	1.00
Hispanic	509	0.17	0.38	0.00	1.00
Multi-race	509	0.11	0.32	0.00	1.00
Teacher/school level					
Attendance hours	23	1.35	2.67	0.00	9.00
PD hours (< 75hs)	23	0.26	0.45	0.00	1.00
PD hours (75-150hs)	23	0.30	0.47	0.00	1.00
PD hours (> 150hs)	23	0.13	0.34	0.00	1.00
Percent of disability	23	26.70	4.17	19.00	30.00
Percent of Minority	23	67.48	22.70	45.00	91.00
Grade 4 science	23	63.83	18.77	38.00	85.00
Grade 8 science	23	32.48	18.21	5.00	49.00

Gender, race, and grade were dummy coded and their differences in student Rasch ability were significant. Results of ANOVA (t-test for gender) were shown in table 5. Female students' average ability was significantly higher than male students ($t = 4.464$, $p < 0.05$), although few male students had the highest ability. Students' ability increased gradually from grade 4 to grade 8 and the differences among groups were significant ($F = 5.691$, $p < 0.001$). The post hoc test illustrated that students in elementary school (grade 4 to 6) shared similar ability and a significant achievement difference was found between elementary school and middle school (grade 7 and 8). Among races, black students scored significantly lower than other races, but there was no significant difference among the rest.

According to the results above, students grade was combined into a new variable that contained elementary and middle schools. Race was recoded into Black, White/Hispanic, and other races that including Native American or Alaska Native, Asian, and multi-races in this case. The recoded control variables were used in the following HLM analyses to increase the accuracy. The overall correlation matrix of student level parameters was shown in Table 6. All the parameters were positively correlated to student ability and thus further analyses were suggested. Furthermore, gender had no correlation with other parameters other than race and a low correlation with outcome. Therefore gender might be removed from further analyses. Finally, none co-linearity was found between the predictors. There is a relatively large correlation between teacher support in science inquiry/discourse and teacher attitude/expectation towards student's work, which might be worth noticing in further analyses.

Table 5 Results of ANOVA (t-test) for control variables

Control Variables	Category	N	Mean Rasch Ability (S.D)	Minimum	Maximum
Gender	Male	246	-.275 (.841)	-2.33	3.31

t = 4.464*	Female	263	-.123 (.789)	-2.33	2.79
Total		509			
Grade	Grade 4	61	-.428 (.689)	-2.23	1.02
F = 5.691***	Grade 5	136	-.324 (.684)	-2.33	1.32
	Grade 6	85	-.304 (.735)	-2.33	1.43
	Grade 7	86	-.102 (.805)	-2.33	1.32
	Grade 8	141	.033 (.972)	-2.33	3.31
Total		509			
Race	Black	157	-.508 (.767)	-2.33	1.32
F = 7.404***	American Indian or Alaska Native	15	-.144 (.583)	-1.04	.91
	White	171	-.085 (.876)	-2.33	3.31
	Hispanic	86	-.081 (.699)	-2.33	1.32
	Multi-race	57	.023 (.775)	-1.46	2.08
	Asian	23	0.952 (.791)	-1.46	1.55
Total		509			

Note: *** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$

Table 6 Correlation matrix of student level parameters

	Ability	Gender	Grade	Race	Self-efficacy	NOS	Teacher support
Ability							
Gender	.093*						
Grade	.196**	.037					
Race	.239**	-.091*	-.042				
Self-efficacy	.198**	-.001	-.082	.061			
NOS	.241**	.000	.111*	.094*	.302**		
Teacher support	.139**	.034	.097*	.057	.244**	.300**	
Teacher attitude	.200**	.050	-.071	.070	.294**	.268**	.460**

Note: *** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$

Results of two-level HLM analyses of student's understanding in disciplinary science

Among student level predictors (Model II), student self-efficacy, understanding of science, and teacher support in discourse & inquiry were positively related to student ability (B = 0.17, $p < 0.01$, B = 0.13, $p < 0.05$, B = 0.16, $p < 0.05$, respectively), when gender, race, and grade were held constant (in Table 7). Students with one point higher in self-efficacy/NOS/teacher support than the grand mean scored 0.17/0.13/0.16 higher in their ability. While teacher attitude/expectation towards student's work was not significant. The variables explained 16% of total variance in student's ability.

The fully unconditional model (Model I) illustrated a significant variance (21%) between classrooms ($u_0 = 0.142$, $p < 0.001$). After adding teacher level predictors of PD and school demographics (Model III), the coefficients of student level variables kept stable. School demographics and teacher PLC attendance were not significantly related to student ability, while there was a significant difference among groups of PD hours since PD was dummy coded and not centered in the model. Students whose teachers were with more than 150 hours PD experience scored significantly higher than those whose teachers with no PD in terms of student ability (B = 0.71, $p < 0.05$). The average student ability difference was 0.71 point between the two groups and favored the group of more PD hours. Furthermore, no difference was found between the three groups of teachers who had PD less than 150 hours by a following ANOVA. Thus, the variable was recombined into two categories that include teacher PD hours more than 150 hours and less than 150 hours. The teacher level parameters explained another 6% of the variance in student's ability and overall 22% variance was explained by the full model.

Table 7 Results of HLM analyses

	Model I	Model II B (s.e.)	Model III B (s.e.)
Student level			
Self efficacy		0.17**(0.05)	0.17**(0.05)
NOS		0.13* (0.05)	0.12* (0.05)
Teacher support		0.16* (0.06)	0.16*(0.06)
Teacher level			
Attendance			-0.06 (0.04)
PD_Hours			
<75			-0.25 (0.17)
75-150			0.02 (0.15)
>150			0.71* (0.25)
u_0	0.142	0.086	0.046
r	0.526	0.476	0.476
Pseudo R2	0.00	0.16	0.22

Note: *** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$

Potential mediation effects

Possible mediation effects between teacher PD hours and student ability were tested through simple regression by using PD hours (> 150) as independent variable and the four significant student-level variables as outcomes. The analyses were conducted by merging teacher PD hours into student level by considering the small size of teachers, which brought a considerable limitation in this step. The recoded variable of PD hours illustrated a significant relationship with student ability (B = 0.51, $p < 0.001$). Students whose teachers were with more than 150 hours PD experience scored 0.51 higher than those whose teachers with less than 150 hours PD. None of the parameters at thestudent level was significantly correlated with teacher PD hours. Nevertheless, student understanding of NOS was marginally significant (B = 0.25, $p < 0.10$), which indicated a possible positive association between them. The difference in student understanding of NOS between the two groups was 0.25 points on average and favored the group with PD hours more than 150 hours. Therefore, a potential mediation effect was proposed between PD hours and student ability through student's understanding of NOS. Further analyses were needed to confirm the effect.

Discussion

PD intervention and teachers' SCK results

Teacher SCK results were not related to total PD hours and summer activity types. The reason might be two-fold. First, the purpose of summer activities was not aligned with test content. In other words, the interdisciplinary science research, curriculum study, and college science course, may not directly relate to teacher subject content knowledge. Second, the category of summer activities might not be specific enough to reveal the differences between groups because a finer separation created too small sample size in subgroups in this case. Heller et al (Heller et al., 2012) studied three types of PD, namely, analyzing teaching cases, looking at student work, and metacognitive analysis. The results showed significant improvement on teacher content knowledge because the coherence between PD and content knowledge. However, none significant relationship between PD and teacher written justification indicated a similar mismatch as this study. The principle of coherence between PD and outcome is essential (Wilson, 2013). Therefore, a consistent assessment of teachers after PD was required.

The PLC session in the PD program was designed towards improving teacher' pedagogy knowledge. However, the attendance in 2012-2013 is not significantly related to teachertest results in summer 2013 while the attendance hours in 2013-2014 became positively correlated with teacher scores in summer 2014. The results indicated that the effect of PLC session might take place in a cumulative fashion by considering pre- and posttest for the same group of teachers in two successive years. The assumption was built on a large overlap rate of teachers (82%) in the two tests. The delay was consistent with previous studies that PD program took a certain time period to show its impact (Garet et al., 2001; Heller et al., 2012; Maerten-Rivera et al., 2015).

PD intervention and student reported teacher classroom practice

Student reported teacher classroom practices included teacher attitude and expectation on student work and teacher support in inquiry and discourse. There was a significant rise in teacher attitude and expectation on student work from 2013 Fall semester to 2014 Spring semester ($F = 11.488$, $p < 0.01$) while teacher support in inquiry and discourse remained the same in two semesters. The simple regression results between PD interventions and the two parameters of teacher classroom practices were not significant, which indicated no relationship. Teacher support in science inquiry/discourse might be difficulty to change compared with their attitude/expectation. With limitation of student self-reported teacher classroom practices, the results were consistent with previous studies of teacher change in classroom practices (Brickhouse, 1990; Garet et al., 2001; Savasci & Berlin, 2012; Windschitl, 2003). However, the reasons of changing practices were not clear in this study and further study is needed to find out possible reasons.

PD intervention and students' ability in interdisciplinary science

The relationships between student ability and student level parameters, which included both control and other survey predictors, were consistent with previous studies. Gender difference (V. E. Lee & Burkam, 1996), racial gap (Catsambis, 1995), and student learning progression (Liu, 2007) in science achievement were significant and they were well controlled to eliminate bias in further analyses. Student reported self-efficacy, understanding of NOS, teacher attitude/expectation on their work, and teacher support in science inquiry/discourse were all positively correlated, which indicated an internal consistency of the survey results. Furthermore, student self-efficacy and understanding of NOS were positively related to their understanding of interdisciplinary science in terms of Rasch ability. The results are consistent with previous studies of student self-efficacy and their science achievement, and provided empirical evidence on effect of NOS (Baker & White, 2003; Lederman, 1992; McComas et al., 2002; Pajares et al., 2000).

Specific research on how teacher expectation/attitude towards student work and teacher support in science inquiry/discourse influenced student science achievement in Elementary/Middle School was rare. A study of lecturer expectation on undergraduate student achievement in bioethics showed positive but weak relations between them (Hanegan et al., 2008). In another study of high school classroom environment, the authors found teacher support, order and organizations, and innovative teaching strategies had positive relationship with student attitudes towards science (Myers & Fouts, 1992). The positive correlations in this study showed similar pattern. By considering the latent trait measured by the instrument, which was understanding of interdisciplinary science, the associations of two teacher parameters were reasonable. Science inquiry and discourse, though difficulty to imply, would improve student conceptual change in science knowledge (Cuevas et al., 2005; Duschl & Osborne, 2002; Keys & Bryan, 2001; Osborne et al., 2004; Wallace & Kang, 2004). The study provided evidence on the positive relationship between teacher support in science inquiry/discourse and student's understanding in disciplinary science.

In HLM analyses, PD hours with 175 hours was the only significant parameter at teacher level and the associations of student and teacher level parameters were independent. The results are consistent with previous studies about duration of effective PD. The time duration of effectiveness varied according to different purpose of the PD (Banilower et al., 2007; Supovitz & Turner, 2000). However, the higher performance of students might also be because of teacher selection process. Teacher beliefs and attitude have been proved to have direct impact on student's achievement as aforementioned. In other words, the teachers who joined longer PD sessions were likely to have more positive attitude and be passionate in learning and developing. Therefore, the influence embedded in their everyday instruction and eventually impacted student achievement. Furthermore, the small sample size of teachers with PD hours more than 150 hours and its sole composition of coordinating teachers may also result in a higher achievement of their students. The coordinate teachers in each school had more experience and responsibility in managing extra duty, however, most background information was not available due to the low response in teacher survey. Further study should use randomized classroom sample to produce a more general results.

How PD eventually related to students ability remained unclear. Through the analyses in the study, student understanding of NOS partially mediated the relationship between PD hours and student ability. Students whose teachers took more than 150 hours PD had higher self-perceived understanding in NOS compared with the other group, and therefore they also had higher ability. While the mediation effects were not found in teacher expectation/attitude and support. The results indicated that the PD program and assessment of student learning outcome might be mismatched. More detailed and specific studies were required to explore the relations in the model of effectiveness of PD aforementioned. Further research design should focus on the alignment of the purpose of the PD program, evaluation of teacher development, reflection of teacher change in knowledge and practices, student change in attitude/beliefs, and assessment of student learning outcome to provide empirical evidence on effectiveness of PD program.

Limitations and Implications

The study had three major limitations. First, the sample teachers were selected by volunteering applications and the sample size was small due to the high missing rate. The small sample size had brought relatively large error in the estimates of teacher level and the overall reliability and validity of analyses might be undermined. Second, the student survey was still under examination of reliability and validity. The parameters generated from the survey, which included student level control variables, predictors, and outcomes, may have possibly over- or under -estimated the relationships involved in the study. Finally, the sample students and teachers in two semesters in the HLM analyses were considered as independent because of a very low rate of overlap. The interpretation of results from the cross-sectional data should be with cautious.

The study provided empirical evidence on effectiveness of PD program, which illustrated that coherence and duration of the program had significant impact. Furthermore, the study shed lights on how effects of PD could finally benefit student learning outcome. The mediation role of student understanding in Nature of Science was only a tiny piece of the map. Further research is called to reveal the relationships among PD program, student learning outcome, student beliefs and attitudes, teacher beliefs, attitudes, and classroom practices. In practice, the project set a good example of partnership between local universities, colleges, and public schools. To improve teacher knowledge and skills by using support and resources from local institutions is mutually beneficial. A stronger bonding between local schools and other institutions is suggested.

Reference

Abd-El-Khalick, F., Bell, R. L., & Lederman, N. G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education, 82*(4), 417-436.

Aitken, G., & Sinnema, C. (2008). Effective pedagogy in social sciences/tikanga ā iwi: Best evidence synthesis iteration.

Alton-Lee, A. (2011). (Using) Evidence for Educational Improvement. *Cambridge Journal of Education, 41*(3), 303-329.

Baker, T. R., & White, S. H. (2003). The effects of GIS on students' attitudes, self-efficacy, and achievement in middle school science classrooms. *Journal of geography, 102*(6), 243-254.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. *Handbook of research on teaching, 4*, 433-456.

Banilower, E. R., Heck, D. J., & Weiss, I. R. (2007). Can professional development make the vision of the standards a reality? The impact of the national science foundation's local systemic change through teacher enhancement initiative. *Journal of Research in Science Teaching, 44*(3), 375-395.

Blank, R. K., de las Alas, N., & Smith, C. (2007). Analysis of the quality of professional development programs for mathematics and science teachers: Findings from a crossstate study. *Washington, DC: Council of Chief State School Officers. Retrieved February, 29,* 2008.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational researcher, 33*(8), 3-15.
Brickhouse, N. W. (1990). Teachers' beliefs about the nature of science and their relationship to classroom practice. *Journal of Teacher Education, 41*(3), 53-62.
Britton, E. D., & Schneider, S. A. (2007). Large-scale assessments in science education. *S. Abell, & N., Lederman, eds. Handbook of research on science education. Mahwah, NJ: Lawrence Erlbaum Associates, Inc*, 1007-1040.
Bybee, R. W. (1995). Science curriculum reform in the United States. *Redesigning the science curriculum*, 12-20.
Bybee, R. W., Carlson-Powell, J., & Trowbridge, L. W. (2008). *Teaching secondary school science: Strategies for developing scientific literacy*: Pearson/Merrill/Prentice Hall.
Cardullo, V. M., & Forsythe, L. (2013). Co-Teaching a New Pedagogical Practice for Pre-Service Teachers. *School-University Partnerships, 6*(2), 90-96.
Catsambis, S. (1995). Gender, race, ethnicity, and science education in the middle grades. *Journal of Research in Science Teaching, 32*(3), 243-257.
Cheng, M. M. H., & So, W. W. M. (2012). Analysing Teacher Professional Development through Professional Dialogue: An Investigation into a University-School Partnership Project on Enquiry Learning. *Journal of Education for Teaching: International Research and Pedagogy, 38*(3), 323-341.
Cobern, W. W., Schuster, D., Adams, B., Skjold, B. A., Muğaloğlu, E. Z., Bentz, A., & Sparks, K. (2014). Pedagogy of Science Teaching Tests: Formative assessments of science teaching orientations. *International Journal of Science Education, 36*(13), 2265-2288.
Cohen, D., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. *The Teachers College Record, 102*(2), 294-343.
Cuevas, P., Lee, O., Hart, J., & Deaktor, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching, 42*(3), 337-357.
Cunningham, C. M. (1998). The effect of teachers' sociological understanding of science (SUS) on curricular innovation. *Research in Science Education, 28*(2), 243-257.
Czerniak, C. M. (2007). Interdisciplinary science teaching. *Handbook of research on science education*, 537-559.
Desimone, L. M. (2009). Improving impact studies of teachers’ professional development: Toward better conceptualizations and measures. *Educational researcher, 38*(3), 181-199.
Diamond, B. S., Maerten-Rivera, J., Rohrer, R., & Lee, O. (2013). Elementary teachers' science content knowledge: Relationships among multiple measures. *Florida Journal of Educational Research, 51*.
Diamond, B. S., Maerten-Rivera, J., Rohrer, R. E., & Lee, O. (2014). Effectiveness of a curricular and professional development intervention at improving elementary teachers' science content knowledge and student achievement outcomes: Year 1 results. *Journal of Research in Science Teaching, 51*(5), 635-658.
Douglas, A. S. (2011). The Different Learning Opportunities Afforded Student Teachers in Four Secondary School Subject Departments in an Initial Teacher

Education School-University Partnership in England. *Journal of Education for Teaching: International Research and Pedagogy, 37*(1), 93-106.
Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education.
Franke, M. L., Carpenter, T. P., Levi, L., & Fennema, E. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal, 38*(3), 653-689.
Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching A Meta-Analysis. *Review of Educational Research, 82*(3), 300-329.
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.
Gess-Newsome, J., & Lederman, N. G. (2001). *Examining pedagogical content knowledge: The construct and its implications for science education* (Vol. 6): Springer Science & Business Media.
Grossman, T., & Hirsch, E. (2009). State Policies to Improve Teacher Professional Development. Issue Brief. *NGA Center for Best Practices*.
Hanegan, N. L., Price, L., & Peterson, J. (2008). Disconnections between Teacher Expectations and Student Confidence in Bioethics. *Science & Education, 17*(8-9), 921-940.
Heller, J. I., Daehler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching, 49*(3), 333-362.
Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.
Klassen, S. (2006). Contextual assessment in science education: Background, issues, and policy. *Science Education, 90*(5), 820-851.
Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' Content Knowledge and Pedagogical Content Knowledge The Role of Structural Differences in Teacher Education. *Journal of Teacher Education, 64*(1), 90-106.
la Velle, L., Reynolds, D., & Nichol, J. (2013). An Unusually Effective School/University Programme: The Plymouth and Peninsula Model. *Journal of Education for Teaching: International Research and Pedagogy, 39*(4), 400-415.
Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*(2), 139-162.
Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*(4), 331-359.
Lee, O., Deaktor, R., Enders, C., & Lambert, J. (2008). Impact of a multiyear professional development intervention on science achievement of culturally and linguistically diverse elementary students. *Journal of Research in Science Teaching, 45*(6), 726-747.

Lee, V. E., & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability level, and course emphasis. *Science Education, 80*(6), 613-650.

Liu, X. (2007). Elementary to High School Students' Growth over an Academic Year in Understanding Concepts of Matter. *Journal of Chemical Education, 84*(11), 1853.

Maerten-Rivera, J. L., Huggins-Manley, A. C., Adamson, K., Lee, O., & Llosa, L. (2015). Development and validation of a measure of elementary teachers' science content knowledge in two multiyear teacher professional development intervention projects. *Journal of Research in Science Teaching, 52*(3), 371-396.

McComas, W. F., Clough, M. P., & Almazroa, H. (2002). The role and character of the nature of science in science education *The nature of science in science education* (pp. 3-39): Springer.

Metz, K. E. (1998). Scientific inquiry within reach of young children. *International handbook of science education, 1*, 81-96.

Mintzes, J. J., Wandersee, J. H., & Novak, J. D. (2005). *Assessing science understanding: A human constructivist view*: Academic Press.

Myers, R. E., & Fouts, J. T. (1992). A cluster analysis of high school science classroom environments and attitude toward science. *Journal of Research in Science Teaching, 29*(9), 929-937.

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of curriculum studies, 19*(4), 317-328.

Nowicki, B. L., Sullivan-Watts, B., Shim, M. K., Young, B., & Pockalny, R. (2013). Factors influencing science content accuracy in elementary inquiry science lessons. *Research in Science Education, 43*(3), 1135-1154.

NRC. (2012). *A Framework for K-12 Science Education:: Practices, Crosscutting Concepts, and Core Ideas*: National Academies Press.

NRC. (2013). Next Generation Science Standards: For States, By States: National Academies Press Washington, DC.

Nuthall, G., & Alton-Lee, A. (1995). Assessing classroom learning: How students use their knowledge and experience to answer classroom achievement test questions in science and social studies. *American Educational Research Journal, 32*(1), 185-223.

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020.

Pajares, F., Britner, S. L., & Valiante, G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology, 25*(4), 406-422.

Pepper, S. K., Hartman, K. J., Blackwell, S. E., & Monroe, A. E. (2012). Creating an Environment of Educational Excellence: The University of Mississippi-PDS Partnership--The Evolution Continues. *School-University Partnerships, 5*(1), 74-88.

Raudenbush, S. W. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*: Scientific Software International.

Roediger III, H. L., & Marsh, E. J. (2005). The positive and negative consequences of multiple-choice testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(5), 1155.

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and

student learning. *Journal of Research in Science Teaching, 48*(2), 117-148. doi:10.1002/tea.20408
Rutherford, F. J., & Ahlgren, A. (1990). *Project 2061: Science for all Americans.* Paper presented at the National Conference of the Dwight D. Eisenhower Mathematics and Science Education Program, Washington, DC.
Savasci, F., & Berlin, D. F. (2012). Science teacher beliefs and classroom practice related to constructivism in different school settings. *Journal of Science Teacher Education, 23*(1), 65-86.
Saxe, G. B., & Gearhart, M. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education, 4*(1), 55-79.
Scher, L., & O'Reilly, F. (2009). Professional Development for K–12 Math and Science Teachers: What Do We Really Know? *Journal of Research on Educational Effectiveness, 2*(3), 209-249.
Seixas, P. (2001). Review of research on social studies. *Handbook of research on teaching, 4*, 558.
Seraphin, K. k. h. e., Philippoff, J., Parisky, A., Degnan, K., & Warren, D. (2013). Teaching Energy Science as Inquiry: Reflections on Professional Development as a Tool to Build Inquiry Teaching Skills for Middle and High School Teachers. *Journal of Science Education & Technology, 22*(3), 235-251.
Shallcross, T., Spink, E., Stephenson, P., & Warwick, P. (2002). How primary trainee teachers perceive the development of their own scientific knowledge: links between confidence, content and competence? *International Journal of Science Education, 24*(12), 1293-1312.
Shroyer, M. G., & Yahnke, S. (2012). Kansas State University Professional Development School Partnership: Improvement for All. *School-University Partnerships, 5*(1), 13-16.
Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher,* 4-14.
Songer, N. B., & Linn, M. C. (1991). How do students' views of science influence knowledge integration. *Journal of Research in Science Teaching, 28*(9), 761-784.
Stolk, M. J., De Jong, O., Bulte, A. M. W., & Pilot, A. (2011). Exploring a Framework for Professional Development in Curriculum Innovation: Empowering Teachers for Designing Context-Based Chemistry Education. *Research in Science Education, 41*(3), 369-388.
Supovitz, J. A., & Turner, H. M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963-980.
Tal, R., & Argaman, S. (2005). Characteristics and Difficulties of Teachers Who Mentor Environmental Inquiry Projects. *Research in Science Education, 35*(4), 363-394.
Trent, J. (2012). Teacher Professional Development through a School-University Partnership. What Role Does Teacher Identity Play? *Australian Journal of Teacher Education, 37*(7).
van Driel, J. H., Meirink, J., Van Veen, K., & Zwart, R. (2012). Current trends and missing links in studies on teacher professional development in science education: a review of design features and quality of research. *Studies in science education, 48*(2), 129-160.

Wallace, C. S., & Kang, N. H. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets. *Journal of Research in Science Teaching, 41*(9), 936-960.

Weinstein, R. S., Gregory, A., & Strambler, M. J. (2004). Intractable self-fulfilling prophecies fifty years after Brown v. Board of Education. *American Psychologist, 59*(6), 511.

Whitehead, J., & Fitzgerald, B. (2006). Professional Learning through a Generative Approach to Mentoring: Lessons from a Training School Partnership and Their Wider Implications. *Journal of Education for Teaching: International Research and Pedagogy, 32*(1), 37-52.

Wilson, S. M. (2013). Professional development for science teachers. *Science, 340*(6130), 310-313.

Windschitl, M. (2003). Inquiry projects in science teacher education: What can investigative experiences reveal about teacher thinking and eventual classroom practice? *Science Education*, *87*(1), 112-143.