V. Serbay Zambak, Daniel M. Alston, Jeff. C. Marshall,
and Andrew M. Tyminski

Convincing Science Teachers for Inquiry-Based Instruction: Guskey's Staff Development Model Revisited

Abstract

For many years, changing beliefs has been considered a prerequisite for changing classroom practices. However, professional development research has also shown that the opposite relationship is also true—change in practice can precede change in beliefs. This study investigated the effect of a one-year professional development program on in-service science teachers' instructional practices, beliefs, and their students' achievement. The professional development program specifically emphasized supporting middle school in-service science teachers to implement inquiry-based instruction. A non-random, single group, interrupted time-series, quasi-experimental design was used to test Guskey's model of staff development and the relationships among changes in teachers' instructional practices, beliefs and their students' achievement. Similar to Guskey's model, findings from this study show the importance that evidence of improved student achievement has on teachers changing their practices. Teacher change models emerging from our data did not find any links between general teacher change variables. However, factors such as teachers' focus on discourse and curriculum were found to be important components of inquiry-based instruction and teachers' contextual beliefs.

Introduction

The concept of inquiry-based instruction (IBI) in science education has been around for approximately half a century (Anderson, 2007). In the past twenty years, organizations such as the National Science Foundation (NSF) have sought to improve the quality of science instruction by initiating reform activities and professional development (PD) programs (Minner, Levy, & Century, 2010). In this time, IBI emerged as one of the prominent instructional methods encouraged within National Research Council policy documents (National Research Council [NRC], 1996; Duschl, Schweingruber, & Shouse, 2007; Schweingruber, Keller, & Quinn, 2011).

Benchmarks for Science Literacy (American Association for the Advancement of Science [AAAS], 1993), *National Science Education Standards* (NRC, 1996; Olson & Loucks-Horsley, 2000), and *Next Generation Science Standards* (NGSS) (Achieve, 2013), all contributed to the science education reform movement and emphasized strategies congruent with IBI. As a result, science teachers are often called to engage in PD programs designed to increase the quality and quantity of this type of instruction. While these programs have been found to change teachers' instructional practices, research is clear that not all teachers benefit in the same way after being involved in a PD program (Little, 1993; Little, 2012). As such, examination, analysis, and evaluation of PD programs are necessary in order to reveal the different aspects that encourage or inhibit their effectiveness.

PD programs encouraging IBI are considered effective to the extent that they change teachers' instructional practices, teacher beliefs, and student achievement (Johnson, 2006; Boyle, Lamprianou, & Boyle, 2005). Therefore, special attention concerning the interaction among these factors is necessary (Avalos, 2011). Additionally, investigating these three factors would enable more insight regarding the continuing development of a teacher change model. Currently the literature is limited in identifying an effective teacher change model, including the relationships among its components (Clarke & Hollingsworth, 2002). Without having a clear understanding of how and why teachers change as a result of a PD program, we as science teacher educators and PD facilitators run the risk of not fully utilizing our time and resources.

This study follows a one-year PD program designed to increase the quantity and quality of middle school in-service science teachers' IBI. We sought to investigate how this PD influenced teachers' instructional practices, teachers' beliefs regarding IBI, and student achievement. In so doing, we set out to further explore the research regarding the key components of Guskey's (2002) model of teacher change. Specifically, we were interested in whether we could add to the research regarding the nature of the sequential relationships of these factors. Our two main research questions for this study were:

1. How is a one-year PD program emphasizing IBI effective in influencing teachers' instructional practices, teacher beliefs, and student achievement?
2. How are teachers' instructional practices, teacher beliefs, and student achievement related in a one-year PD program emphasizing IBI?

Theoretical Framework

Guskey's staff development model (2002) describes PD as an attempt to

Keywords: teacher change, inquiry based instruction, professional development, teacher beliefs

change teachers' instructional practices, to develop beliefs and attitudes towards different teaching approaches, and to improve students' learning outcomes. The traditional teacher change model suggests that a change in teacher beliefs and attitudes precedes a change in classroom practice. Guskey (2002) alternatively proposed a different model where changes in teachers' instructional practices and student achievement precede changes in psychological factors (e.g., beliefs).

Because of the components emphasized in our PD (e.g., inquiry-based instructional practices, teacher beliefs, and student achievement), we developed an adapted version of Guskey's teacher change model to align with our goals (Figure 1).

Each one of these components was operationalized for this study (Appendix). For example, we defined teachers' instructional practices in terms of four factors (i.e., instruction, discourse, assessment, and curriculum). We hypothesized for this PD that teachers' beliefs could change when they observe that there is an improvement in students' achievement as a result of their IBI practices. In addition, we also predicted that teacher change would occur in a sequential but non-linear fashion.

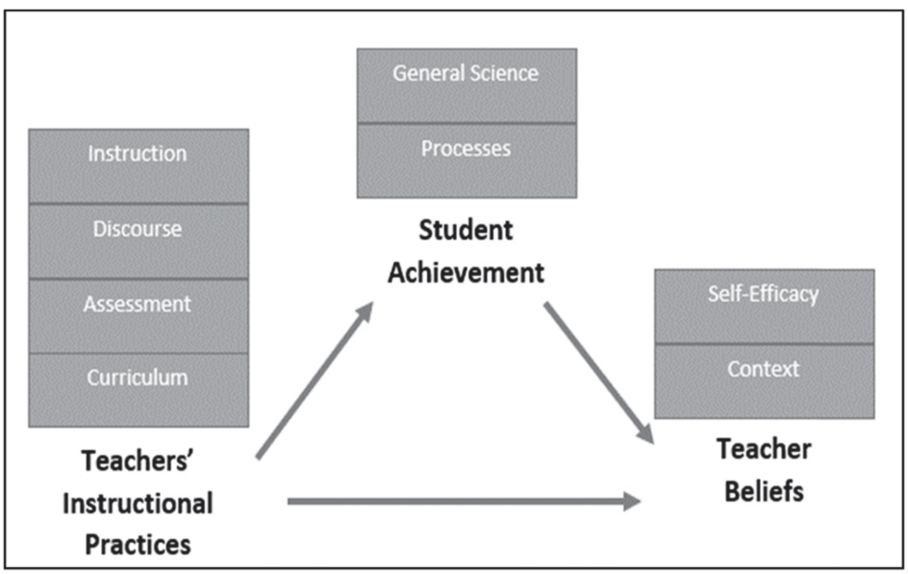

Figure 1. Teacher Change Model for this Study

Literature Review

Inquiry-Based Instruction

Since its introduction over 50 years ago (Anderson, 2007), IBI has grown in prevalence as an effective teaching strategy in science education. While this concept is oftentimes misinterpreted, recent reform documents (e.g., Achieve, 2013; Schweingruber, Keller, & Quinn, 2011) now use the term scientific practices instead of inquiry "to better specify what is meant by inquiry in science and the range of cognitive, social, and physical practices that it requires" (p. 30). With this distinction, we are seeking to make clear that doing any hands-on activity or experiment is not sufficient to be considered science. Students need to engage in the practices used by scientists to appreciate "how scientists establish credibility for the claims that they advance" (Osborne, 2014, p. 180). With these ideas in mind, IBI can be defined as any intentional student-centered instruction that designs experiences that provide students the opportunity to deepen their understanding of scientific content and formulate an accurate conception of the process undergone by scientists to find out and validate knowledge. The following sections provide a brief overview of the effectiveness of PD programs, their impact on teacher beliefs, student achievement, and teachers' instructional practices in the context of IBI.

Effectiveness of Professional Development Programs

Increased accountability for schools and teachers has encouraged stakeholders in science education to search for ways of increasing student achievement (Parise & Spillane, 2010). As a result, policy makers, school administrators, and district leaders pay special attention to the effectiveness of PD programs in terms of the effect they have on student achievement. There are mixed results concerning the impact that PDs emphasizing IBI have on student achievement (Minner, Levy, & Century, 2010). However, among these mixed results, positive relationships are recorded illustrating that PDs encouraging IBI can increase student performance (Minner, Levy, & Century, 2010). Furthermore, researchers have found certain characteristics of PDs to have a positive influence on student achievement. Of these characteristics, teacher reflection and metacognition, as well as collaboration between teachers, administrators and researchers were found to be effective at increasing students' achievement (Heller, Daehler, Wong, Shinohara, & Miratrix, 2012).

PDs having a specific pedagogical and content focus have been found to have positive results in terms of influencing teachers' instructional practices. Researchers indicated engaging teachers in active learning impacts their instructional strategies, where active learning was defined as teachers' deeper understanding of student learning (Desimone, Porter, Garet, Yoon, & Birman, 2002). Moreover, they found PDs focused on a set of higher order instructional practices yield substantial benefits. Higher order instructional methods are described as opportunities for learners to be engaged with active and inquiry-based learning (Raizen, 1998). Similarly, Buczynski and Hansen (2010) used survey findings to demonstrate evidence of teachers' changing their instructional practices towards IBI. Participating teachers were taking traditional science courses from university professors. They also attended workshops that utilized hands-on experiences and demonstrations of IBI. Ninety-two percent of teachers reported they increased the frequency of inquiry-based learning activities when they returned to their classrooms during their instructional practices. Admittedly, teachers' self-reports may represent a skewed reality. However, the cited studies do provide evidence that PDs emphasizing IBI can positively impact teacher instructional practice. Changes in teachers' instructional practices are often tied with the beliefs that teachers hold (Jones & Carter, 2007). Therefore, it is crucial to consider how teacher beliefs are related to their choice of engaging in IBI.

Changing Teachers' Beliefs

Teachers' decisions to change their instructional practices as encouraged by educational reformers are rooted in their beliefs about the nature of science and teaching (Borko & Shavelson, 1990; Anderson, 2002). Clark and Peterson (1986) state that science teachers' beliefs may play a major role in the effectiveness of science education reforms since teachers' beliefs lead to actions and therefore impact students' achievement.

When considering beliefs, researchers have differentiated between the types of beliefs that can impact behavior. Efficacy beliefs and context beliefs are two influential types of beliefs described in the literature that guide teachers' practice. While there are other types of beliefs delineated in the literature regarding educators' teaching practices (e.g., Fang, 1996; Ertmer, 2005), these are outside the scope of the current research study. *Self-efficacy beliefs* have been defined as "beliefs in one's capabilities to organize and execute the course of action required to produce given attainments" (Bandura, 1996, p.3). It has been found that teachers with high self-efficacy beliefs are more likely to implement new teaching practices (Guskey, 1988; Allinder, 1994; Ghaith & Yaghi, 1997). While self-efficacy beliefs impact teacher practice, Ashton (1985) posits that without careful attention paid to context beliefs, the effects of changing efficacy beliefs will be short lived. *Context beliefs* are "those beliefs about the ability of external factors or people to enable a person to reach a goal plus the belief that a factor is likely to occur" (Lumpe & Chambers, 2001, p. 95). Milner, Sondergeld, Demir, Johnson, and Czerniak (2012) showed that teacher beliefs were influenced by the people with whom they worked (e.g., coworkers and administration). In other words, teachers who have high context beliefs (e.g., support from their administrators and coworkers) have a greater probability of attempting to introduce new instructional practices.

Every PD program implements its activities and grounds its philosophical stance according to the teacher change model it adopts (Guskey, 2002). Therefore, a test of theory on teacher change models is required to examine the relationship between teachers' instructional practices, their beliefs about IBI, and student achievement (Supovitz & Turner, 2000).

Methods

A non-random, single group, interrupted time-series, quasi-experimental design (Shadish, Cook, & Campbell, 2002) was used to investigate the research questions. Participants were voluntarily involved in the PD, and their instructional practices, beliefs and students' achievement were measured at various points in time. This was an interrupted time series design because the measurement of participants' instructional practices was also captured prior to the start of the PD.

Participants and the Intervention

Seventy middle school science teachers from a southeastern state were recruited to be part of a PD program designed to improve the quality of IBI. Teachers were provided two weeks of teacher training in the summer, four whole group follow-up sessions during the academic year, four or more full-class observations with debriefing after, and numerous individual support sessions. The emphasis for the summer portion of the PD program and the follow-up meetings was to provide the teachers with modeled IBI lessons, ensure opportunities to reflect on IBI lessons, and collaborate to design inquiry-based lessons that could be implemented.

The planning model that we encouraged teachers to use in our PD program was the 4Ex2 Instructional Model (Marshall, Horton, & Smart, 2009), an adaptation of the 5E Instructional Model (Bybee et al., 2006). The 4Ex2 Model uses the following instructional sequence: Engage, Explore, Explain, and Extend. The "x2" stands for the continuous formative assessment and teacher reflection that should occur throughout each of the four stages.

Five cohorts were used for this study. While many of the participants were involved in the program more than one year, only the first year data were used in this study. Forty science teachers who were teaching in 10 different middle schools became the participants for this study. Table 1 details the gender, ethnicity, and teaching experience distribution of the participating teachers.

These teachers saw IBI as an instruction that provides students with a chance to explore scientific concepts and take responsibility for their learning. Further, teachers largely saw that they should be acting as a facilitator when implementing IBI.

Instruments and Data Collection

In an effort to capture the relationships among teachers' belief changes, instructional practice changes, and student

Table 1. Demographic Information for Participants

Gender	Ethnicity	Teaching Experience	Grade
Male: 7 (17.5%)	Black: 5 (12.5%)	Less than 10 Years: 19 (47.5%)	6th Grade Teacher: 17 (42.5%)
Female: 33 (82.5%)	White: 35 (87.5%)	11-20 Years: 9 (22.5%)	7th Grade Teacher: 13 (32.5%)
		21-30 Years: 10 (25%)	8th Grade Teacher: 10 (25%)
		More than 30 Years: 2 (5%)	

achievement, we analyzed the results from a belief survey, the data from classroom observations measured using the Electronic Quality of Inquiry Protocol (EQUIP) (Marshall, 2009; Marshall, Horton, & White, 2009; Marshall, Smart, & Horton, 2010), and the achievement scores of students measured by the Measures of Academic Progress (MAP) test.

In the belief survey, participants responded to questions about *their self-efficacy* pertaining to the use of IBI methods and about *their context* pertaining to the support they received from their colleagues and administration. The belief survey contained 17 Likert-scale questions and was administered at the beginning and the end of the year-long PD. Each participant filled out the survey online through Survey Monkey. The reliability coefficient for the belief survey was $\alpha = .54$. Context beliefs and self-efficacy beliefs were two sub-constructs measured within the belief survey (Marshall, Horton, Igo, & Switzer, 2009). Context beliefs were measured with a two-item subscale, and self-efficacy beliefs were measured with a four-item subscale. Cronbach's alpha for each subscale was .87. For the comparison of participants' beliefs pertaining to IBI before and after the intervention and for the analysis of the teacher change model, we used the overall score for the belief survey as well as scores for its sub-constructs.

The EQUIP (Marshall, 2009; Marshall, Horton, & White, 2009; Marshall, Smart, & Horton, 2010) was used to score each participant's instructional practices relative to IBI. The EQUIP measures the quality of IBI in regards to four components (i.e., instruction, assessment, discourse, curriculum) and provides a score for each of these constructs as well as an overall inquiry lesson score (Marshall, & Horton, 2011). Scores range from level 1 (pre-inquiry) to 4 (exemplary inquiry). Marshall and Horton (2011) measured Cronbach alpha to be .92, .93, .85, and .81 for each construct respectively, and .94 for the overall lesson score. The reliability coefficient for the entire instrument was .98. Researchers trained on the EQUIP observed participants' classrooms once prior to the PD, and at least once each 9 weeks during the academic year. Depending on individual participant's needs, there occasionally was more than one observation within a 9-week period.

Student achievement was measured by the MAP test (Northwest Evaluation Association, 2000). MAP is a computer adaptive test that assesses students' academic progress with respect to national and state curricula. The science test has two portions that measure general science process skills and content. Therefore, the test is an effective measure of how students are growing in their knowledge of science, as well as their knowledge of science practices. There were 4,496 students of participating teachers who took the science MAP test. The mean for the number of students per teacher was 112 (SD = 44).

Students took the test twice, once early in the fall term and again in the mid- to late-spring term. The reports for each test, prepared by the Northwest Evaluation Association, included raw mean scores for all students. Each teacher had an average score for his/her students' general science content and process skills as well as an overall average composite MAP score. During the final follow-up meeting, participating teachers had a chance to view students' progress from fall to the spring semester. Since teachers took the post-survey after this follow-up meeting, it was feasible for us to examine the link between changes in their beliefs compared to the evidence of student achievement.

Data Analysis

Quantitative data from belief surveys, MAP tests, and the EQUIP were compiled in an Excel spreadsheet. Any teacher who was missing overall data from one of the main data resources was omitted from the analysis. For participants who had an incomplete data entry from one of the data sources, their missing scores were replaced with the average response for that specific question—the mean substitution method (Roth, 1994).

For pre- and post-belief surveys, each participant had an overall sum score, and sum scores for two sub-constructs (e.g., self-efficacy and context) of the survey. The belief change was measured for each participant by computing the difference between sum scores from post- and pre-surveys.

EQUIP observation scores were averaged when multiple observations were collected during a given 9-week period. In order to see the trend for participants' instructional practices, we computed the averages for the pre-PD observations, the first semester observations (i.e., the first two observations), the second semester observations (i.e., the last two observations), and the average of four observations within the academic year. The change in teachers' instructional practices was measured by computing the difference between these observational averages. The same measurements were conducted for each construct of the observation protocol.

The first round of analyses focused on the investigation of the overall change in instructional practices and beliefs of participants after a year of PD. We also examined changes in each of the four EQUIP constructs. Furthermore, we examined students' overall achievement growth (i.e., composite score) from fall to spring, as well as their specific growth in general science content and process skills. To compare each of the scores mentioned, we conducted dependent sample t-test analyses and repeated measures ANOVA on the data using SPSS. During the second round of analyses, we conducted several multiple regression analyses using SPSS in order to test the relationships between instructional practices, student achievement, and teacher beliefs.

Findings and Discussion

Research Question 1: Effectiveness of the PD Program

There was a significant increase ($\bar{X}_{pre} = 77.46$, $\bar{X}_{post} = 81.25$, $p < .05$) in the mean of participants' overall belief about IBI from pre- to post-survey. The program also significantly increased their self-efficacy concerning the use of IBI in their classes

Table 2. Repeated Measures ANOVA Summary for Changes in Teachers' Instructional Practices

	SS	df	MS	F	P
Overall Score	996.715	1	996.715	1253.927	.000
Composite Score pertaining to Instruction	1161.620	1	1161.620	1156.283	.000
Composite Score pertaining to Discourse	937.445	1	937.445	983.996	.000
Composite Score pertaining to Assessment	1026.045	1	1026.045	1131.827	.000
Composite Score pertaining to Curriculum	947.866	1	947.866	1900.195	.000

($\bar{X}_{pre} = 18.25,\ \bar{X}_{post} = 19.60,\ p = 0.03$). After a year involved in the program, teachers' context beliefs also increased. While this increase was not significant (p = .51), it is notable.

The F-Ratio from Repeated Measures ANOVA was significant (p < .001) for the overall observation score and for each composite score (N = 40) for the five observations (Table 2). To understand how participants' instructional practices changed throughout the academic year, a dependent sample t-test was conducted; and a plot was constructed to illustrate this change. Figure 2 indicates that the overall observation and composite scores for sub-constructs before our PD program were lower than the scores during the program. The dependent sample t-test revealed that the overall observation and individual construct scores before the program started were significantly lower than the observation scores recorded throughout the year (p < .05 for each). Figure 2 also shows that participating teachers' overall mean EQUIP score was highest during the second observation, which was taken at the end of the first semester. This result indicates that participating teachers demonstrated the required practices for IBI at the end of the first semester, but these changes in their instructional practices were not sustainable during the second semester.

Repeated Measures ANOVA on MAP scores in the fall and the spring (Table 3) indicated a significant difference (p < .001) for the overall composite score and for general science content and process standards separately. However, considering the previous finding on teachers' sustainability of their reformed classroom practice during the second semester, it is difficult to link these findings to teachers' IBI as teachers did not consistently use IBI in their classrooms throughout the year. In order to determine if there was a relationship, we examined the results from the multiple linear regression analysis.

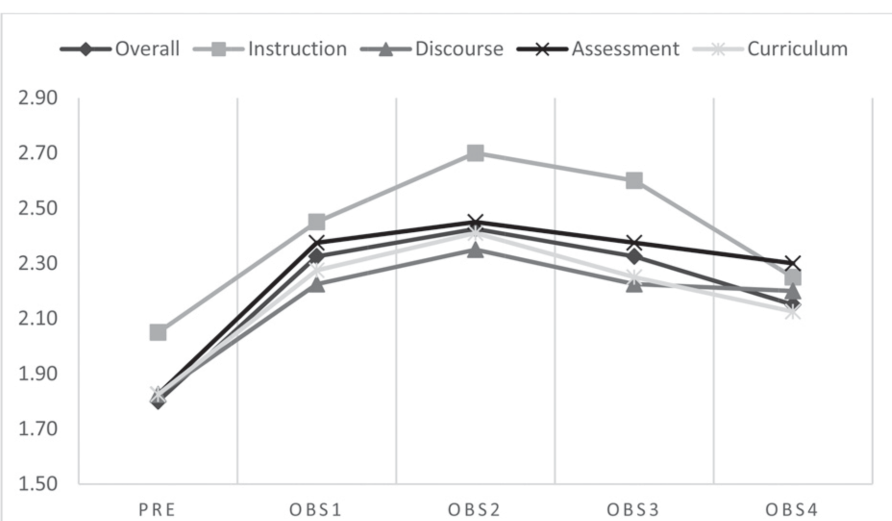

Figure 2. Changes in Instructional Practices among Five Observations

Research Question 2: Teacher Change Process

The relationship among participants' belief change, instructional practice change, and their students' achievement growth was identified through multiple linear regression analysis. In order to explore the teacher change process within one year and how the variables mentioned were interconnected, we computed the power and the effect size of each of the multiple regression findings (Table 4).

Teachers' belief change, as a dependent variable, was not a significant predictor. However, we discovered that change in participants' context belief was predicted by the change in teachers' instructional practice with a low to medium Cohen (1988) effect size. R^2 for this finding is not very strong, which suggests that other significant variables are missing from the equation. β for the predicting variable was negative, which pointed out the inverse relationship between the dependent and predicting variables.

While more research is needed to determine the reason for this inverse relationship, it can be interpreted that as teachers engage in more IBI they become more aware of the support they need from curriculum and administration. Therefore, they realize what they believed to be available actually was not and their context beliefs decrease. Teachers' perception about the support provided in their environment to implement reformed practices could be different. Moreover, they become increasingly aware of the need for a higher level of support within the school as they try to incorporate IBI in their classes.

Teachers' change in their instructional practices with respect to the use of discourse was significantly predicted by students' growth in both general science content and process skills from fall to spring (R^2 = .123, ES = .14). In other words, students' growth in their general science content and inquiry skills in one year contributed to the teachers' questioning, communication pattern and classroom interactions among students. While students' achievement

Table 3. Repeated Measures ANOVA Summary for MAP Scores

	SS	df	MS	F	P
Overall Map Score	3280521.400	1	3280521.400	66203.038	.000
Content Score	3359766.043	1	3359766.043	57676.454	.000
Process Score	3318248.623	1	3318248.623	62867.340	.000

growth predicted teachers' change in the discourse component, it did not predict their change in other components of IBI measured by the EQUIP.

The relationship between teachers' change in their instructional practices and evidence of students' achievement corresponds to Guskey's model (2002). Teachers in this PD seemed to change their IBI practices as long as they saw evidence of growth in their students' achievement, but the change in their teaching practices only occurred in the discourse component of IBI. Teachers' focus on the discourse techniques as a legitimate dimension for the adoption of IBI in their classes might imply that teachers need more than one year within a PD in order to effectively integrate all components of IBI measured by the EQUIP.

Whereas teachers' change in their instructional practices was predicted by their students' achievement, the last multiple regression equation also pointed out the existence of a relationship in the other direction. Students' growth from fall to spring was dependent on teachers' change in their teaching practices with respect to the discourse and curriculum constructs ($R^2 = .254$, ES = .34). As long as teachers used more questioning techniques and discourse through guided and prescribed activities rather than letting students discover by themselves, students improved in general science content and process skills. Rather than minimal guidance, strong guidance (Kirschner, Sweller, & Clark, 2006) enhanced with questioning techniques seems to be the general IBI interpretation of teachers.

Conclusion

Teacher change is a complex process and is therefore hard to describe with a few variables and their relationships. This difficulty comes, in part, because of the fact that every teacher engages in this process differently depending on their beliefs, experiences, needs, and expectations. This PD experience was effective in enabling teachers' to change their beliefs and their instructional practices. It contributed to the achievement of these teachers' students in general science content knowledge and process skills (now denoted as the scientific practices) in one year. Teachers significantly changed their context and self-efficacy beliefs towards the implementation of IBI. Moreover, they changed their instructional practices throughout the year, with the most significant change occurring at the middle of the year.

Significant increase of the mean value for context beliefs from pre- to post-survey showed evidence that support from colleagues, administration, and teacher educators was important for these teachers during the change process. Although their teaching was better than their baseline, the majority of teachers reported that they increased their self-efficacy within one year. However, synthesizing the results from the dependent t-test for self-efficacy and the multiple regression findings (negative coefficient) indicates that teachers' self-efficacy might be a naïve conception which results from the teachers' "awareness of and commitment to constructivism among educators" rather than from their actual classroom practices (Pomeroy, 1993, p. 272).

This study also reinforces the importance of including evidence of student achievement in any models supporting teacher change in practice. Having periodic tests might allow teachers to evaluate how much their changes in instructional practices contribute to their students' academic achievement. In other words, teacher educators should emphasize showing evidence of students' achievement in order to convince teachers to change their practices. Convincing teachers in terms of student learning outcomes may then initiate a change in their beliefs about reform and IBI.

Even though the data collected could not show evidence supporting the entirety of Guskey's model (2002), this PD program emphasizing IBI provided some insights regarding the links among the major variables within the teacher change process. The model emerging from the data did not indicate any links between variables in general, but they were evident when analyzing the sub-constructs. Regarding the results here, a modified teacher change model can be drawn as shown in Figure 3, which presents a reciprocal relationship between the change in teachers' instructional practices and their students' achievement, and a one-way directional relationship between teachers' change in their instructional practices and their beliefs.

The strength of the claims made by this study is somewhat limited in terms of teacher change process and the sequence of proper teacher change through a PD program. For stronger

Table 4. Findings on Multiple Regression Analysis

Dependent Variable	Predictors	R^2	Significance	Standardized ß	Power	Effect Size
Changes in Teacher Beliefs (Context)	Changes in Teacher Instructional Practices	.121	.035	-1.45	.62	0.14
Changes in Teacher Instructional Practices (Discourse)	Changes in Student MAP Scores	.123	.031	.351	.64	0.14
Changes in Student MAP Scores	Changes in Teacher Instructional Practices (Discourse and Curriculum)	.254	.019	.599 -.434	.89	0.34

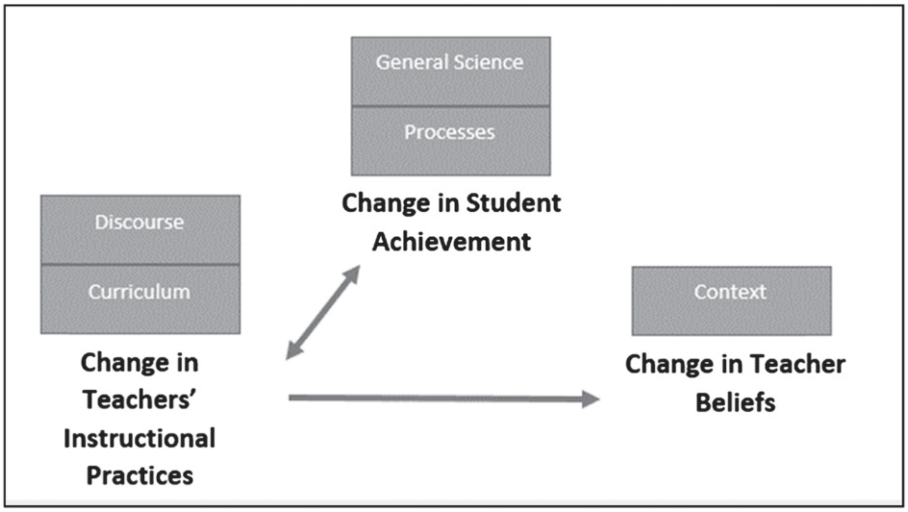

Figure 3. Teacher Change Model for IBI

claims about the relationships between teachers' change in their instructional practices and their beliefs as well as their students' achievement, there is a need for a larger sample so that statistical methods such as Structural Equation Modeling or Hierarchical Linear Modeling might be applied to the data. Secondly, the data only represented teachers' change for one-year. Relationships among variables and the sequence might be different for longer PD experiences. Further research should look at the trends of larger samples involved in a PD program for more than one-year.

References

Achieve. (2013). *Next generation science standards for today's students and tomorrow's workforce.* Washington, DC: Achieve. Available at www.nextgenscience.org

Allinder, R. M. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education*, *17*(2), 86-95.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. Washington, DC: ERIC Clearinghouse.

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry? *Journal of Science Teacher Education*, *13*(1), 1-12.

Anderson, R. D. (2007). Inquiry as an organizing theme for science curricula. In S. K. Abell, & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 807-828). New Jersey: Lawrence Erlbaum Associates, Inc.

Ashton, P. (1985). Motivation and the teachers' sense of efficacy. In C. Ames, & R. Ames (Eds.), *Research on motivation in education* (pp. 141-167). Orlando, FL: Academic Press.

Avalos, B. (2011). Teacher professional development in Teaching and Teacher Education over ten years. *Teaching and Teacher Education*, *27*(1), 10-20.

Bandura, A. (1996). *Self-efficacy: The exercise of control*. New York: Freeman.

Borko, H., & Shavelson, R. (1990). Teacher decision making. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 311–346). London, UK: Routledge.

Boyle, B., Lamprianou, I., & Boyle, T. (2005). A longitudinal study of teacher change: what makes professional development effective? Report of the second year of the study, school effectiveness and school improvement. *An International Journal of Research, Policy and Practice*, *16*(1), 1-27.

Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education*, *26*(3), 599-607.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins and effectiveness.* Colorado Springs: BSCS.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255-296). New York: Macmillan.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, *18*(8), 947-967.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Lawrence Erlbaum.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81-112.

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration? *Educational Technology Research and Development*, *53*(4), 25-39.

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research*, *38*(1), 47-65.

Ghaith, G., & Yaghi, H. (1997). Relationships among experience, teacher efficacy and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, *13*(4), 451-458.

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes toward the implementation of instructional innovation. *Teaching and Teacher Education*, *4*(1), 63-69.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice*, *8*(3), 381-391.

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.

Johnson, C. C. (2006). Effective professional development and change in practice: Barriers science teachers encounter and implications for reform. *School Science and Mathematics*, *106*(3), 150-161.

Jones, M. G., & Carter, G. (2007). Science Teacher Attitudes and Beliefs. In S. K. Abell, & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 1067-1104). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75-86.

Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, *15*(2), 129-151.

Little, J. W. (2012). Professional community and professional development in the learning-centered school. In M. Koy & K. van Veen (Eds.), *Teacher learning that matters: International perspectives* (pp. 22-46). New York, NY: Routledge.

Lumpe, A. T., & Chambers, E. (2001). Assessing Teachers' Context Beliefs. *Journal of Research on Technology in Education, 34*(1), 93-107.

Marshall, J. C. (2009). *The creation, validation, and reliability associated with the EQUIP (Electronic Quality of Inquiry Protocol): A measure of inquiry-based instruction*. Paper presented at the National Association of Researchers of Science Teaching Conference.

Marshall, J. C., & Horton, R. M. (2011). The Relationship of Teacher-Facilitated, Inquiry-Based Instruction to Student Higher-Order Thinking. *School Science and Mathematics*, *111*(3), 93-101.

Marshall, J. C., Horton, R., Igo, B. L., & Switzer, D. M. (2009). K-12 science and mathematics teachers' beliefs about and use of inquiry in the classroom. *International Journal of Science and Mathematics Education*, *7*(3), 575-596.

Marshall, J. C., Horton, B., & Smart, J. (2009). 4E× 2 instructional model: Uniting three learning constructs to improve praxis in science and mathematics classrooms. *Journal of Science Teacher Education*, *20*(6), 501-516.

Marshall, J. C., Horton, B., & White, C. (2009). EQUIPping teachers: A protocol to guide and improve inquiry-based instruction. *The Science Teacher*, *76*(4), 46–53.

Marshall, J. C., Smart, J., & Horton, R. M. (2010). The design and validation of EQUIP: An instrument to assess inquiry-based instruction. *International Journal of Science and Mathematics Education*, *8*(2), 299-321.

Milner, A. R., Sondergeld, T. A., Demir, A., Johnson, C. C., & Czerniak, C. M. (2012). Elementary teachers' beliefs about teaching science and classroom practice: An examination of pre/post NCLB testing in science. *Journal of Science Teacher Education*, *23*(2), 111-132.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, *47*(4), 474-496.

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

National Research Council (Ed.). (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. The National Academies.

Northwest Evaluation Association. (2000). *Measures of academic progress*. Portland, OR: Author.

Olson, S., & Loucks-Horsley, S. (Eds.). (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.

Osborne, J. (2014). Teaching scientific practices: Meeting the challenge of change. *Journal of Science Teacher Education*, *25*(2), 177-196.

Parise, L. M., & Spillane, J. P. (2010). Teacher learning and instructional change: How formal and on-the-job learning opportunities predict change in elementary school teachers' practice. *The Elementary School Journal*, *110*(3), 323-346.

Pomeroy, D. (1993). Implications of teachers' beliefs about the nature of science: Comparison of the beliefs of scientists, secondary science teachers, and elementary teachers. *Science Education*, *77*(3), 261–278.

Raizen, S. A. (1998). Standards for science education. *Teachers College Record*, *100*(1), 66-121.

Roth, P. L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology*, *47*(3), 537-560.

Schweingruber, H., Keller, T., & Quinn, H. (Eds.). (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

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.

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

V. Serbay Zambak, PhD, Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, WI. Correspondence concerning this article should be addressed to: V. Serbay Zambak, PhD, Post-Doc Research Associate, Marquette University, Department of MSCS, Cudahy Hall 337, Milwaukee, WI 53201-1881. Email: vecihi.zambak@marquette.edu

Daniel M. Alston, PhD, Assistant Professor, Department of Reading and Elementary Education, College of Education, University of North Carolina-Charlotte, NC.

Jeff. C. Marshall, PhD, Professor of Science Education and Chair, Department of Teaching and Learning, College of Education, Clemson University, Clemson, SC.

Andrew M. Tyminski, PhD, Associate Professor, Department of Teaching and Learning, College of Education, Clemson University, SC 29634, Clemson, USA

Appendix

Main Components	Sub-components	Operational Definition
Teacher Instructional Practices	Instruction	Instructional strategies; the teacher's and students' role during the instruction
	Discourse	The level and complexity of the teacher's questioning; classroom interactions among students and the teacher
	Assessment	Assessment of students' prior knowledge; conceptual development and student reflection.
	Curriculum	The depth of the content and standards covered; the organization of the lesson.
Student Achievement	General Science	Students' understanding of concepts related to life, physical, earth and space sciences.
	Concepts and Process	Students' content skills to analyze, synthesize key ideas in science, and process skills to do scientific inquiry.
Teacher Beliefs	Self-Efficacy	Teachers' confidence while operating on a specific activity
	Contextual	Teachers' perception of the availability of support from their colleagues and administrators to implement IBI.