International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

The impact of a professional development model on middle school science teachers' efficacy and implementation of inquiry

Christine Lotter, Whitney Smiley, Stephen Thompson & Tammiee Dickenson

To cite this article: Christine Lotter, Whitney Smiley, Stephen Thompson & Tammiee Dickenson (2016) The impact of a professional development model on middle school science teachers' efficacy and implementation of inquiry, International Journal of Science Education, 38:18, 2712-2741, DOI: 10.1080/09500693.2016.1259535

To link to this article: https://doi.org/10.1080/09500693.2016.1259535

Published online: 20 Dec 2016.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2016
VOL. 38, NO. 18, 2712–2741
http://dx.doi.org/10.1080/09500693.2016.1259535

Routledge
Taylor & Francis Group

The impact of a professional development model on middle school science teachers' efficacy and implementation of inquiry

Christine Lotter[a], Whitney Smiley[b], Stephen Thompson[a] and Tammiee Dickenson[b]

[a]College of Education, Instruction & Teacher Education, University of South Carolina, Columbia, SC, USA; [b]Office of Program Evaluation, College of Education, University of South Carolina, Columbia, SC, USA

ABSTRACT

This study investigated a professional development model designed to improve teachers' inquiry teaching efficacy as well as the quality of their inquiry instruction through engaging teachers in practice-teaching and reflection sessions. The programme began with a two-week summer Institute focused on both inquiry pedagogy and science content and continued with academic year support for participants' inquiry implementation. An inquiry teaching efficacy instrument was administered 3 times to 25 teacher participants to gauge changes in their personal self-efficacy and outcome expectancy across 5 essential features of classroom inquiry. To examine actual practices, pre/post classroom observations of the teachers' inquiry enactments were evaluated using a quality of inquiry observation protocol. Following the summer Institute, teachers had statistically significant increases in their self-efficacy for teaching inquiry in four of the five essential features and increases in one of the five essential features for outcome expectancy. Teachers' quality of inquiry teaching also increased after the professional development programme. We discuss implications of this PD model for moving teachers towards implementation of new instructional techniques as well as the influence of a supportive school community on teachers' efficacy with inquiry instruction.

ARTICLE HISTORY
Received 28 March 2016
Accepted 7 November 2016

KEYWORDS
Inquiry; professional development; self-efficacy; teacher beliefs

For over two decades, K-12 science education reform in the United States has emphasised student engagement in inquiry and scientific practices (National Research Council (NRC), 2012; *NSES*, NRC, 1996). Guiding reform documents such as the National Science Education Standards for Science Education (NRC, 1996) defined a set of five essential features of student inquiry including engagement with questions, collection and analysis of data stemming from investigations, explanation of concepts, and communication of results (NRC, 2000). These and other science practices such as modelling and argumentation were further elaborated and connected directly to content expectations in the *Next Generation Science Standards* (NGSS Lead States, 2013). Despite this emphasis on engaging students in science practices and inquiry

CONTACT Christine Lotter lotter@mailbox.sc.edu College of Education, Instruction & Teacher Education, University of South Carolina, 820 Main Street, Columbia, SC 29208, USA

investigations, researchers have described a lack of student engagement in inquiry within U.S. secondary science classrooms (Capps & Crawford, 2013a; Windschitl, 2002b). Related research highlights that teacher beliefs about science teaching inhibit the growth of such instructional practices within U.S. science classrooms (Lotter, Harwood, & Bonner, 2007; Luft, 2001; Magnusson, Krajick, & Borko, 1999; Marco-Bujosa, McNeill, González-Howard, & Loper, 2016). Specifically, teachers' self-efficacy, or their beliefs about their ability to teach science effectively, has been shown to influence the effort they expend and the skill with which they implement new instructional strategies (Enochs, Scharmann, & Riggs, 1995; Lakshmanan, Heath, Perlmutter, & Elder, 2011; Ross & Bruce, 2007). Thus, efforts to move U.S. teachers forward with the teaching practices that are emphasised in the *Next Generation Science Standards* (NGSS Lead States, 2013) require an investigation into the effectiveness of professional development (PD) models aimed at changing teacher practice as well as an investigation into how teacher beliefs and actions influence these PD outcomes.

Research into science teacher PD reveals multiple programme components that seem to impact teacher beliefs and augment the instructional gains experienced by science teachers who participate in a given PD experience. For example, engaging teachers in experiences that demonstrate a given methodology has been identified as an important factor impacting their understanding of the instructional processes required to enact it (O'Brien, 1992). Likewise, providing teachers with authentic experiences that engage them in the skills and knowledge necessary for effective science teaching has been shown to positively impact teacher beliefs and practices (Brown, Collins, & DuGuid, 1989; Capps, Crawford, & Constas, 2012; Greeno, 1997). Similarly, social construction of new knowledge and practices through reflective dialogue and interaction with other educators has been shown to enhance teacher beliefs and practices (Akerson, Cullen, & Hanson, 2009; Lave & Wenger, 1991; Lotter, Yow, & Peters, 2014; Loughran, 2014). Other PD components such as content-focused instruction, engaging teachers in developing unit plans using a new strategy (Capps et al., 2012), and opportunities to co-teach with experienced peers or content-specific coaches (Vogt & Rogalla, 2009; Zwart, Wubbels, Bolhuis, & Bergen, 2008) have been identified as being important to teacher growth as well. In addition to the components of the PD, the support teachers receive from administrators, colleagues, and PD providers affects how reform practices are enacted in teacher's classrooms (Roehrig, Kruse, & Kern, 2007; Weis, 2013). Thus, this study sought to investigate how one PD model, which included many of the key components advocated for by current PD literature, impacted participating science teachers' beliefs and practices about inquiry-based instruction.

Literature review

To support our study, we describe current literature on teacher learning, the theoretical foundations of our PD model, and how teacher learning is influenced by teacher beliefs, including teacher efficacy beliefs. With a focus on inquiry science teaching, we also describe this construct and research surrounding inquiry PD.

Teacher learning

The PD model advocated for in this research views the 'science teacher as learner' which implies that teachers are viewed as active participants in their learning; shaping their learning outcomes rather than being passive recipients of knowledge (Loughran, 2014, p. 811). With this PD model, teacher learning is viewed as context-dependent and socially constructed through dialogue and collaboration with others (Putnam & Borko, 2000). Furthermore, one of the major theories supporting our PD efforts, situated cognition, emphasises how learning is enhanced through participation in authentic experiences that provide opportunities for teachers to develop the skills and knowledge necessary for effective teaching (Brown et al., 1989; Greeno, 1997).

Our PD model also seeks to engage teachers in communities of practice (Lave & Wenger, 1991) in which teachers socially construct new knowledge and practices through reflective dialogue and interaction with other educators. Such communities have been shown to enhance teacher learning in other PD projects as well as our own previous research that has supported the creation of our current PD model (Akerson et al., 2009; Lotter, Yow, & Peters, 2014; Loughran, 2014). Similarly, Akerson et al. (2009) developed a PD programme using a community of practice framework that improved elementary teachers' nature of science instruction through collaborative discussions and problem-solving. Teacher reflection on new practices has been shown to be vital to the building of an effective community of practice that builds teacher knowledge and skills (Capps & Crawford, 2013b; Kolb, 1984). For example, Marco-Bujosa et al. (2016) found that middle school teachers who engaged in active reflection while implementing a reform-based educative curriculum showed greater learning gains compared to teachers who just viewed the curriculum as a source of activities. Capps and Crawford (2013b) also found that teacher reflection was an important component to changing teachers' beliefs during an inquiry and nature of science PD.

The continuation of this reflective community of practice into a teacher's school through administrative and peer approval of new instructional strategies as well as support for teacher collaboration and teacher change is also important for teacher learning (Schriver & Czerniak, 1999). Related to this, in their large-scale study of PD in Australia, Ingvarson, Meiers, and Beavis (2005) found that teachers' knowledge and practice are

> … enhanced by the extent to which that program also strengthens the level of professional community in the school; that is, the extent to which it increases opportunities for teachers to talk about the specifics of their teaching practice and student learning, share ideas and support each other as they attempt to implement ideas from the professional development program. (p. 14)

Similarly, other studies have shown school support to be an important element in effective PD (Roehrig et al., 2007; Schriver & Czerniak, 1999; Tschannen-Moran & Hoy, 2007) and teachers' effective science instruction (Weis, 2013). Related to this, school-level science administrator support has been found to be an important form of support that influences teachers' enactment of inquiry-based chemistry curriculum (Roehrig et al., 2007). As Ingvarson et al. (2005) state that in their PD evaluation study 'the level of school support comes through the analysis as an important enabling condition with a significant shaping influence on the opportunities to learn that teachers experience' (p. 15).

Therefore, our study hypothesised that school support would also be important for teachers' learning and use of inquiry-based teaching strategies.

In addition to utilising communities of practice, our PD model also seeks to create teacher cognitive dissonance, as this has been thought to influence teacher growth during PD experiences. Such PD experiences challenge teachers' existing conceptions and/or create cognitive dissonance in teachers in order to change their pre-existing beliefs and practices (Loughran, 2014; Thompson & Zeuli, 1999). Some have argued that teachers experiencing cognitive dissonance within a PD is more important for changing teacher practice than increasing teacher self-efficacy (Feldman, 2000; Gregoire, 2003; Settlage, Southerland, Smith, & Ceglie, 2009; Wheatley, 2002). The model of PD presented in this study incorporates the components discussed here within a yearlong PD experience for middle school teachers. However, the model of PD presented in this study rests on the notion that providing these PD experiences alone is not sufficient. We view teacher beliefs as being connected with their teaching practices and our PD model highlights how experience, coupled with reflection on experience, may lead to change in beliefs (Fishman, Marx, Best, & Revital, 2003; Richardson, 1996).

Self-efficacy

The beliefs a teacher holds about teaching and learning have a direct influence on the teachers' instructional practices (Fang, 1996), including attempts to implement inquiry-based instruction (Capps et al., 2012; Luft, 2001; Magnusson et al., 1999). For example, Luft (2001) found that beliefs differed between induction and experienced teachers and these beliefs influenced how teachers enacted extended inquiry cycles after a PD programme.

Other researchers assert that self-efficacy beliefs related to classroom practices are the most important factor influencing teaching practices (Pajares, 1996). Pajares (1996) defined self-efficacy as an 'expectancy belief' or an 'individuals' perceived capabilities to attain designated types of performances and achieve specific results' (p. 546). Two dimensions of efficacy beliefs are described by Bandura's (1986) social cognitive theory: personal self-efficacy and outcome expectancy. Bandura (1997) defines personal self-efficacy as 'a judgment of one's ability to organize and execute given types of performances, whereas an outcome expectation is a judgment of the likely consequence such performance will produce' (p. 21). Efficacy beliefs are thought to be task- and context-specific (Tschannen-Moran, Hoy, & Hoy, 1998). For example, a teacher might have a high perceived efficacy for teaching chemistry but a low efficacy for teaching physics or for teaching physics using inquiry strategies.

In addition to individual teacher efficacy, Bandura (1997) defined collective teacher efficacy as a groups' 'shared belief in its conjoint capabilities to organize and execute courses of action required to produce given levels of attainments' (p. 477). Groups of teachers, typically those experiencing similar pressures and goals, may experience similar changes in their efficacy through their interaction (Goddard, Hoy, & Hoy, 2000). Collective efficacy is often associated with teachers at successful schools (Klassen, Tze, Betts, & Gordon, 2010), but may also apply to effective PD that has teachers work together as a community to improve student learning. Goddard et al. (2000) described two key features to the development of collective teacher efficacy: teachers' analysis of their teaching task and their assessment of their teaching competence relative to the rest of the group (p. 485).

Teaching efficacy seems to impact teacher growth. For example, teachers with higher teaching self-efficacy have been shown to experiment more with new instructional techniques and often exert more effort to succeed even when teaching tasks are perceived as difficult (Enochs et al., 1995; Ross & Bruce, 2007; Tschannen-Moran et al., 1998). Similarly, Lakshmanan et al. (2011) found that teachers who went into their PD programme with higher science teaching outcome expectancies maintained a higher level of reform teaching than teachers with lower initial science teaching outcome expectancies. Interestingly, they found statistically significant growth in the teachers' science teaching efficacy, but not their outcome expectancy. Likewise, Sinclair, Naizer, and Ledbetter (2011) found significant increases in K-8 teachers' science teaching self-efficacy and outcome expectancy after a PD that focused on science content and inquiry instruction with confirmation of effective classroom practice using a classroom observation instrument.

Inquiry teaching

In *A Framework for K-12 Science Education* (NRC, 2012), U.S. educators and scientists outlined a vision for the next generation of science standards. Within this vision is an emphasis on U.S. students learning content through scientific practices that include students in asking questions, developing and using models, planning and carrying out investigations, analysing and interpreting data, constructing explanations, engaging in argument from evidence, and obtaining, evaluating, and communicating information (pp. 3–5). These science practices align well with the conception of inquiry teaching outlined in the previous *National Science Education Standards* (NRC, 1996). Inquiry teaching consists of a variety of instructional strategies that teachers use to guide students to an understanding of the science content standards and science process skills (Anderson, 2007). The *NSES* (NRC, 2000) described five features as essential for students when participating in inquiry: (1) 'engaging in scientifically oriented questions', (2) 'giving priority to evidence in responding to questions', (3) 'formulating explanations from evidence', (4) 'connecting explanations to scientific knowledge', and (5) 'communicating and justifying explanations' (NRC, 2000, p. 29). Because our programme took place before the adoption of the *NGSS* (NGSS Lead States, 2013), these five essential features of inquiry were taught explicitly to teachers during our PD programme. The five essential features are also a component of the inquiry efficacy belief instrument used in this study (Smolleck, Zembal-Saul, & Yoder, 2006).

The nature of inquiry teaching presents pedagogical challenges for teachers. Inquiry teaching is featured as a way to develop a rich understanding of science content as well as a process that can promote learners' understanding of scientific reasoning strategies and knowledge generation (NRC, 1999, 2012). Few teachers have experienced teaching or learning science as inquiry (Blanchard, Southerland, & Granger, 2009; Kleine et al., 2002; Windschitl, 2002a), and research suggests that the model of science that teachers *have* experienced may negatively impact their attempts to implement inquiry teaching (Amanti, 2000; Krajcik, Blumenfeld, Marx, & Soloway, 2000; Metz, 2000). For example, a key facet of inquiry learning and teaching is dealing with data. The processes of analysing, interpreting, and arguing from data are components of knowledge generation central to scientific inquiry. It is through engagement in these same processes that learners come to understand important aspects of scientific reasoning. However, as Krajcik et al., (2000) highlighted, many science teachers have no experience with these aspects of inquiry and are more likely to have dealt with data from structured laboratory experiments where a predetermined answer was known. In a more recent study, Capps and Crawford (2013a) found that only 6 of 26 experienced elementary and middle-level teachers engaged their students in inquiry investigations in their classrooms. Examples like these support the call for science teachers to develop a more accurate conceptual framework for characterising good inquiry teaching (White & Frederiksen, 2000). Such a conceptual framework becomes a cognitive tool that allows teachers to reflect on teaching practices as a prerequisite for enhanced inquiry teaching (White & Frederiksen, 2000). Development of such a framework may help ease the 'reality shock' often encountered by teachers who experience and attempt to implement new instructional strategies (Tschannen-Moran et al., 1998).

Research questions

Given this literature review and the need for additional research on inquiry-based PD (Capps et al., 2012), our study set out to investigate the following research questions:

(1) How does an inquiry PD model with cycles of practice-teaching and reflection influence teachers' perceived self-efficacy and outcome expectancy for teaching science as inquiry?
(2) How does an inquiry PD model with cycles of practice-teaching and reflection influence teachers' quality of inquiry instruction?
(3) What teacher and school factors impact teachers' efficacy and implementation of inquiry science teaching?
(4) What is the relationship between teachers' initial efficacy for teaching science as inquiry and their inquiry implementation?

PD model

The study reported here provided teachers with an inquiry teaching conceptual framework through a PD model designed to engage teachers in inquiry content instruction, practice-teaching to students, and collaborative reflection on inquiry teaching. The PD experience began with a two-week summer institute (Institute) and continued with three follow-up sessions during the academic year. The Institute was divided into 4 main segments over the 2-week period (7 hours a day for 10 days): whole-group inquiry pedagogy instruction, small group content instruction, practice-teaching with middle school students, and small group reflection sessions.

Inquiry pedagogy sessions

During the Institute, the teachers spent from 30 to 60 minutes engaged in a morning session in which they participated in inquiry-based activities and discussions of inquiry, questioning, argumentation, and scientific modelling. In each morning session, an education faculty or experienced educator engaged the participants in a different pedagogical skill (e.g. argumentation, scientific modelling) through an interactive activity that modelled this pedagogy. The sessions were accompanied by presentations and/or readings about the theories underlying the targeted pedagogical strategy, thereby helping teachers to understand the instructional processes required to enact it (O'Brien, 1992).

Content instruction

After the whole-group pedagogy session, teachers were divided into three content area teams (energy, genetics, and astronomy) by grade level. In each group, a University content instructor (Ph.D. in science) as well as a pedagogical instructor (middle school teacher who previously completed this PD) led the teachers through inquiry lessons. The science content was taught using locally developed middle school project-based curriculum units. These project-based units engaged students in inquiry investigations in which they gathered and analysed data to explain scientific concepts. For example, the sixth-grade Energy unit used a driving question of 'How can I build a house that will keep me cool in our state?' to help the teachers investigate the state standards around energy and electricity. As part of this unit, teachers built and tested 'coolers' with different insulating or conducting materials, analysing temperature data to determine which materials would keep a liquid (water) warm for the longest time. As a culminating project, they combined all their energy and electricity content knowledge to design an efficient house that would keep them cool in 90-degree summer heat. During these sessions, the teachers learned content by experiencing a middle school curriculum as students and then delving deeper into the content through discussion about the experiences with the content instructors.

In both the pedagogy and content sessions, teachers were taught and practised two structured protocols to improve students' scientific explanations: Claim, Evidence, and Reasoning (McNeil & Krajcik, 2008) and Predict, Observe, and Explain (POE, White & Gunstone, 1992). As identified above, student's ability to write scientific explanations is one of the essential features of inquiry instruction (NRC, 2000). In a melting block activity in the six-grade content session, teachers were asked to predict which block (metal, plastic, or wood) would melt an ice cube the fastest. The teachers had to write and share aloud both their prediction and an explanation for their thinking before investigating (observe phase) the results. After observing the results and taking additional temperature data, teachers then had to record and explain their findings using the Claim, Evidence, Reasoning framework (McNeil & Krajcik, 2008). Teachers wrote their claim or initial answer, described their evidence (data and interpretations), and then provided reasoning that connected their evidence to scientific theory. This framework was used throughout the pedagogy sessions as well as during the content sessions.

Practice-teaching

During the first two days of the Institute, the content instructors taught lessons from the curriculum to groups ($n = 30$) of middle school students in a summer enrichment programme. Students were recruited from local schools for a Science and Technology Enrichment Program (two 90-minute science classes each day). After observing this model teaching for 2 days, teacher teams (3–5 teachers) practice-taught lessons adapted from the curriculum used during the content sessions to the students. This practice-teaching allowed participants to immediately enact their newly learnt inquiry strategies and work through the new curriculum in a low-stakes environment. Each teacher team led six 90-minute classes to small groups of 8–15 students. The groups engaged in collaborative planning for the teaching experiences, and engaged in daily reflection (as described in the following section). Collectively, these opportunities provided teachers with authentic experiences that engaged them in the skills and knowledge necessary for effective teaching, which has been shown to positively impact teacher beliefs and practices (Brown et al., 1989; Capps et al., 2012; Greeno, 1997).

Reflection on teaching

All practice-teaching sessions were observed by at least one project staff member (e.g. model teacher or science instructor). After the practice-teaching sessions, project staff met with the teachers for a one-hour collaborative reflection discussion centred on their strengths and 'missed opportunities' during the practice inquiry teaching. Within these reflective groups, the teachers set instructional goals to improve aspects of their teaching before the next practice-teaching session. Project staff helped the teachers to set attainable goals that resulted in improving the teachers' practice and increasing their efficacy for inquiry. For example, a teacher might set a goal to integrate more opportunities for students to engage in argument from evidence during the next teaching day. During reflection sessions, project staff focused on 'missed opportunities' instead of weaknesses and praised teachers' positive movement towards reaching their goals to encourage positive affective states (Singer, Lotter, Gates, & Feller, 2011). These reflective teaching groups became communities of practice (Lave & Wenger, 1991) that promoted teachers' construction of new knowledge and practices through reflective dialogue and interaction with other educators (Akerson et al., 2009; Lotter, Yow, & Peters, 2014; Loughran, 2014).

Academic year follow-up

The PD model continued during the academic year with three 4-hour Saturday workshops held at a local middle school. These workshops began with an hour-long whole-group pedagogy session focused on topics such as questioning, formative assessment, and student-driven content explanations. Then, the participants returned to their same content groups from the Institute for two hours of new grade-specific inquiry lessons designed to increase their content knowledge. Teachers also spent an hour sharing and reflecting on their current use of inquiry-based practices in their classrooms. In addition to the workshops, the project staff observed teachers and provided in-classroom support as requested by the teachers. In-classroom support consisted of a staff member observing a teacher's enactment of inquiry methods and providing feedback to the teacher during a post-observation conference. Teachers also had access to instructional materials needed to enact many of the inquiry lessons in the form of a lending library (science probeware, computers, etc.). Collectively, the academic year components were designed to mitigate known obstacles to inquiry-based teaching and provide continued support for teachers who were attempting to implement these newly learned strategies within their home teaching contexts.

Methods

Participants

Middle school teachers from the Southeastern United States participating in a one-year inquiry-based PD programme were the subjects of this study. This study describes results from the first year of a larger three-year state-funded grant project that is described further elsewhere (Lotter, Thompson, Dickenson, Smiley, Bilue, & Rea, in press). A total of 25 teachers completed all research requirements (3 efficacy surveys, pre-institute and post-institute observation requirements, and written reflections) and were included in this study. Although 48 teachers participated in the PD and all but 5 completed all 3 administrations of the survey, only 30 of these teachers allowed researchers to observe or record a final inquiry lesson in their classrooms. Although this is a limitation of this study, the researchers are confident that the 25 participating teachers are representative of the larger population of middle school teachers from this region based on comparison of teacher EQUIP scores across multiple grant years and similarities in teacher survey responses across all 48 participants. Demographic information on the participating teachers (highest level of education, years of teaching experience) and schools (percentage of students receiving free and reduced lunch) can be found in Table 1. Teachers voluntarily applied for the programme and project staff chose applicants based on a first-come first-served basis with priority given to teachers from high-need school districts.

Data sources and analysis

This study used a mixed-methods approach, specifically parallel mixed analysis (Tashakkori & Teddlie, 1998), to gather information and make sense of teachers' perceived efficacy

Table 1. Demographic information for participating middle-level teachers.

Pseudonym	Grade taught	Years teaching experience	% Free and reduced lunch of teacher's school	Highest level of education (BA = Bachelor of Arts, BS = Bachelor of Science M = Masters)
Carrie	6	8	26	BA
Betty	6	26	58	M
Leslie	6	22	26	BS
Pam	6	14	66	M
Angie	6	5	62	BA
Jessica	6	1	86	BS
Sue	6	11	91	BS
Lindsey	7	26	77	M
Jaime	7	13	42	M
Mary	7	7	78	BA
Lori	7	3.5	91	M
Ed	7	4	48	M
Caitlin	7	29	77	BA
Laura	7	11	76	M
Julie	7	7	40	M
George	8	4	67	BS
Mark	8	7	26	BS
Christina	8	7	18	BS
Tina	8	8	18	BS
Patty	8	15	45	M
Diane	8	3	36	M
Melissa	8	27	26	BS
Courtney	8	2	76	BS
Marie	8	25	52	BS
Connie	8	11	78	M
Average		11.86	55.44	

for Teaching Science as Inquiry (TSI) and their actual inquiry teaching methods. With the dual purpose of measuring teacher change over time and understanding the complex phenomena of teachers' use of inquiry after a PD programme, a combination of qualitative and quantitative research strategies was used (Newman, Ridenour, Newman, & DeMarco, 2003). To measure changes in teacher efficacy, teachers completed the TSI (Smolleck et al., 2006) efficacy instrument before and after the summer institute and at the end of the academic year (EOY). To evaluate teachers' classroom enactment of inquiry, we observed and evaluated one inquiry lesson before the Institute and one from the school year following the Institute using the Electronic Quality of Inquiry Protocol (EQUIP, Marshall, Smart, & Horton, 2010). To investigate how the teachers' beliefs about the PD structure may have influenced their efficacy to teach through inquiry, we collected and coded teachers' responses to open-ended questionnaires given before and after the two-week Institute and again at the EOY. At the end of the Institute, we also collected a final written reflection from each teacher that required feedback related to the various PD components. Through parallel mixed analysis, both quantitative and qualitative data sources were collected and analysed concurrently to triangulate and provide additional depth of information to support or refute the quantitative efficacy data (Tashakkori & Teddlie, 1998).

Teaching science as inquiry

The TSI instrument (Smolleck et al., 2006) was used to assess the teachers' personal self-efficacy and outcome expectancy for TSI. The TSI includes 34 items developed as indicators of teachers' personal self-efficacy to implement inquiry strategies and 35 items developed as indicators of outcome expectancy for students (Smolleck et al., 2006; Smolleck & Yoder, 2008). As described in Smolleck et al. (2006), the TSI items were cross-classified to reflect one of the five essential features of inquiry as described by *NSES* (NRC, 2000). For example, 'I have the necessary skills to determine the best manner through which children can obtain scientific evidence' was categorised as an item that measured teacher self-efficacy related to Essential Feature 3 (Dira-Smolleck, 2004, p. 2). Similarly, 'My students will make use of data in order to develop explanations as a result of teacher guidance' was categorised as an item that measured outcome expectancy related to Essential Feature 3 (Dira-Smolleck, 2004, p. 3). Table 2 shows the distribution of TSI items by inquiry feature and efficacy type. Although the TSI was originally developed for use with elementary teachers, the items are aligned to the five essential features of inquiry that are applicable across K-12 education (NRC, 2000). The directions to our survey asked the teachers to think about their current classroom teaching when answering each item.

Table 2. Distribution of TSI items by inquiry feature efficacy type.

Essential features of inquiry (NRC, 2000)	No of personal self-efficacy Items	No of outcome expectancy items	No of items per inquiry feature
Learner engages in scientifically oriented questions	7	8	15
Learner gives priority to evidence in responding to questions	8	8	16
Learner formulates explanations from evidence	6	7	13
Learner connects explanations to scientific knowledge	6	4	10
Learner communicates and justifies explanations	7	8	15
Totals	34	35	69

The TSI instrument was administered to participants both prior to and following the Institute, and again at the end of the academic school year (March–April). Each of the items used a five-point Likert-type response scale that ranged from *Strongly Disagree* to *Strongly Agree* and used a midpoint of *Uncertain*. Reliability analysis using Cronbach alpha indicated a reliability of .90 or higher for all scales with our sample. Further validity and reliability information is described in Smolleck et al. (2006) and Smolleck and Yoder (2008).

Quantitative analysis

We examined the TSI at three different time points from each teacher. Due to having multiple measurement occasions from each participant, the assumption of independence of observations was not met, and thus a statistical method that allowed correlations among observations was needed. Typically in these types of longitudinal studies, repeated-measure Analysis of Variance (ANOVA) is used to detect differences between means. However, in this case, multilevel growth modelling was used where time (lower level) was nested within teachers (higher level), which allows for measurement occasions from the same participant to be correlated. Additionally, this technique offers the flexibility to model covariates at the teacher level, thus allowing us to model how teacher demographic variables impact growth. In addition, if we have a balanced data set and use Restricted Maximum Likelihood estimation, the usual analysis of variances based on *F*-tests and *t*-tests can be derived from the multilevel regression results (Raudenbush, 1993). Thus, the repeated-measure ANOVA is a special case of the more general multilevel regression model, with the multilevel regression model allowing more flexibility when modelling.

Pre- and post-institute inquiry lesson analysis

Participants video-recorded or had project staff observe one inquiry lesson before and after PD participation. Teachers chose the lessons they wanted to have observed based on their individual definition of inquiry. Each lesson was analysed using the EQUIP (Marshall et al., 2010), a valid and reliable instrument that measures science teachers' quality of inquiry teaching. The EQUIP consists of 19 Indicators divided into 4 factors to evaluate inquiry-based instruction: (1) Instruction, (2) Discourse, (3) Assessment, and (4) Curriculum. Researchers also determine a holistic score for each factor based on the overall level of inquiry in the lesson (e.g. Total Instruction score) and a Total Inquiry level for the lesson. For each factor, inquiry is divided into Pre-Inquiry (Level 1), Developing Inquiry (Level 2), Proficient Inquiry (Level 3), and Exemplary Inquiry (Level 4). Each level of inquiry is operationally defined for each factor. According to Marshall et al. (2010), Level 3 or Proficient Inquiry is aligned with the level of inquiry in current national standards documents.

Prior to this study, a science education researcher and two programme staff members (former middle and high school science teachers) were trained on the EQUIP through a set of training videos with expert scores provided on the instrument website. The training continued until all raters came to at least 95% agreement without comparing scores. All inquiry lessons in this study were scored by one of these previously trained researchers.

Questionnaires and written reflections

The teachers completed three questionnaires online using SurveyMonkey® with short answer questions pertaining to their understanding of inquiry, their beliefs about student learning, and the impact of the programme on their instruction. The questionnaire also included a set of eight questions on a five-point Likert scale from Strongly Agree to Strongly Disagree that were averaged to create a School Support score. The School Support questions were adapted from the Teacher Opinion and Preparedness section of the 2006 Horizon Research Local Systemic Change through Teacher Enhancement Science (6–12) Questionnaire and included items such as, 'I feel supported by colleagues to try out new ideas in teaching science' and 'I have time during the regular school week to work with my peers on science curriculum and instruction' (Horizon Research, 2006). The pre-institute questionnaire was completed in May before the Institute, the post-institute questionnaire was completed on the last day of the Institute, and the end-of-year questionnaire was completed the following March. The teachers also completed a written reflection on the last day of the Institute in which they commented on the value of each of the four main components of the Institute (pedagogy, content, practice-teaching, and reflection sessions).

Analysis of surveys and reflections

The surveys and reflection papers were read multiple times and patterns related to the teachers' conceptions of inquiry and the value of the Institute components were coded using a constant comparative method (Bogdan & Biklen, 1998). For example, teachers' comments illustrating how programme components helped them to gain confidence with inquiry (confirming evidence), as well as evidence showing any negative influence of workshop components (disconfirming evidence), were coded. We also coded the survey and reflection data for general themes relating to the teachers' discussion of instructional strategies, their definitions of inquiry teaching, and discussion of student learning. For example, teachers' definitions of inquiry were reduced into themes and organised into tables based on the survey administration as well as the teacher's initial efficacy level (high or low initial efficacy). After separating the responses into general themes (inquiry definitions, questioning, science practices, etc.), teacher statements within these themes were further categorised. For example, teachers' responses around questioning were further separated into statements addressing teacher behaviours (teachers asking thinking questions, not answering their own questions), student behaviours (providing experiences to inspire student questions), and specific questioning strategies (e.g. wait time). These themes are summarised in both the narrative and in Table 10 (inquiry categories) in the results section.

Results

Efficacy for TSI

Teachers' mean self-efficacy and outcome expectancy (with standard deviations in parentheses) for TSI as measured by the TSI increased from the pre-Institute ($M = 132.76$ (14.37) and 139.28 (17.55)) to post-Institute ($M = 142.64$ (17.10) and 141.88 (21.98)) administrations, but fell slightly from post-Institute to the follow-up ($M = 137.68$ (17.63) and 140.28 (18.99)) administration at the EOY. Figure 1 graphically illustrates the teachers' total self-efficacy and outcome expectancy scores as a per cent of the total scales. Per cents were calculated due to the scales containing unequal numbers of items. The self-efficacy scale has a maximum value of 165 and the outcome expectancy scale has a maximum value of 175. The teachers held high self-efficacy and outcome efficacy beliefs about TSI even before participating in the institute with scores on all administrations ranging between 79% and 86% of the total scale points.

We examined the TSI at three different time points controlling for several demographic factors and therefore used a multilevel growth model with time points nested within individuals. As described above, the TSI has two major subscales, each with five essential areas of inquiry (Table 2), thus the results for this instrument are broken down into the two sections across the five essential inquiry features.

Self-efficacy

Teachers' self-efficacy rose in four out of five essential features from pre-Institute to post-Institute, but fell slightly at the end-of-the-year administration. However, the overall change from pre-Institute to end of year remained positive. The means, mean differences, and standard deviations for the self-efficacy scale across all essential features of inquiry are presented in Table 3.

Additionally, the multilevel growth model results are presented in Table 4, with **bolded** estimates indicating statistical significance at the 0.05 level. In this table, the intercept shows the overall mean for each of the five essential features of inquiry for the pre-Institute controlling for gender, years teaching, free/reduced lunch, and school support. Teachers self-efficacy for four out of the five essential inquiry features were impacted by their perceived

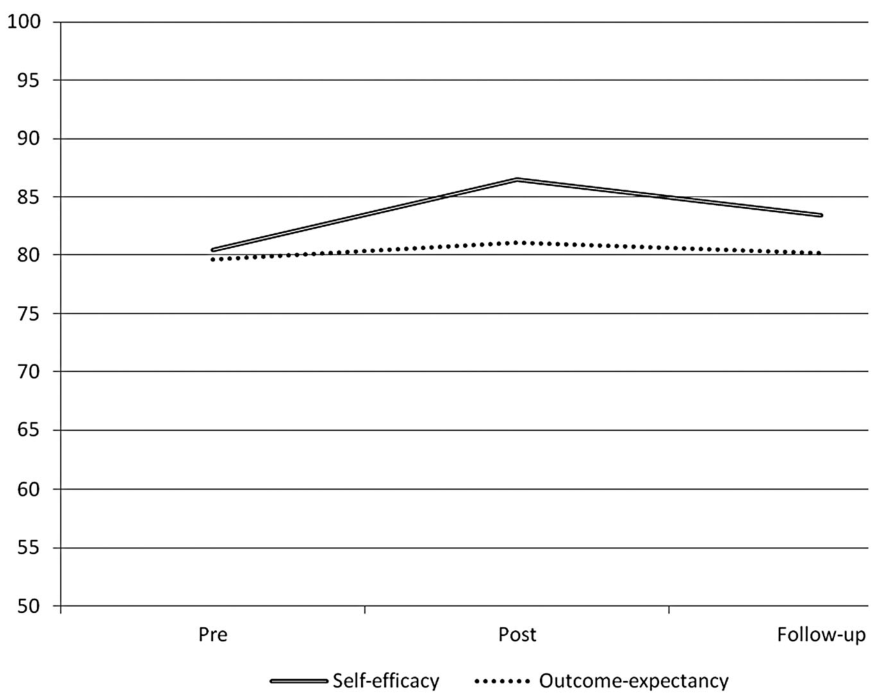

Figure 1. Average per cent of the total self-efficacy and outcome expectancy scales across time.

Table 3. Mean values and mean differences across essential areas of inquiry for self-efficacy.

Inquiry feature	Mean (SD) pre	Mean (SD) post	Mean (SD) follow-up	Mean difference post-pre (SD)	Mean difference follow-post (SD)	Mean difference follow-pre (SD)
Learner engages in scientifically oriented questions	27.8 (3.39)	29.72 (4.62)	28.8 (4.72)	1.92 (4.79)	−0.92 (3.41)	1.00 (5.09)
Learner gives priority to evidence in responding to questions	31.16 (4.18)	32.04 (4.82)	31.6 (5.26)	0.88 (4.31)	−0.44 (4.63)	0.44 (3.83)
Learner formulates explanations from evidence	20.52 (2.33)	22.4 (2.75)	21.56 (2.71)	1.88 (2.88)	−0.84 (2.69)	1.04 (2.68)
Learner connects explanations to scientific knowledge	24.2 (2.97)	27.28 (2.61)	25.84 (2.91)	3.08 (2.56)	−1.44 (2.50)	1.64 (3.19)
Learner communicates and justifies explanations	29.08 (3.16)	31.2 (3.98)	29.88 (3.71)	2.12 (3.82)	−1.32 (3.68)	0.80 (3.27)

school support, such that teachers who reported more school support also had higher self-efficacy scores. No significant relationships were found with teacher self-efficacy and the teacher's gender, years of teaching, or school percentage of free/reduced lunch.

In Table 4, the value of Time 2 shows the increase from pre-Institute to post-Institute while the value for Time 3 shows the increase from pre-Institute to the one-year follow-up. For example, for *Learner engages in scientifically oriented questions*, the average increase from pre-Institute to post-Institute (Time 2) was 2.22 points while the average increase from pre-Institute to one-year follow-up (Time 3) was 1.04, indicating that scores increased at post-Institute and declined from post-Institute to the one-year follow-up, but were still higher than the pre-Institute. As indicated in Table 4, three inquiry features had statistically significant increases from pre to post only (*learner engages in scientifically oriented questions, learner formulates explanations from evidence*, *learner communicates and justifies explanations*) with average increases ranging from 2.04 points 2.35 points. Teachers' self-efficacy towards the inquiry feature, *learner connects explanations to scientific knowledge*, showed significant changes across all three administrations, with scores increasing an average of

Table 4. Multilevel modelling results for self-efficacy in each essential feature of scientific inquiry.

	Learner engages in scientifically oriented questions		Learner gives priority to evidence in responding to questions		Learner formulates explanations from evidence		Learner connects explanations to scientific knowledge		Learner communicates and justifies explanations	
	Coeff.	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*
Intercept	**23.97**	**4.93**	**25.7**	**5.63**	**19.2**	**3.21**	**18.89**	**3.27**	**25.21**	**4.12**
Time 2	**2.22**	**0.96**	1.22	0.88	**2.04**	**0.57**	**3.17**	**0.59**	**2.35**	**0.77**
Time 3	1.04	0.96	0.57	0.88	1.09	0.57	**1.74**	**0.59**	0.87	0.77
Gender	−2.02	2.09	−2.37	2.39	−1.46	1.36	−1.69	1.39	−3.03	1.75
Years teaching	−0.09	0.08	−0.15	0.09	−0.04	0.05	−0.05	0.05	−0.12	0.07
FRL	−0.04	0.03	−0.04	0.03	−0.03	0.02	−0.02	0.02	−0.03	0.03
School support	**2.61**	**1.47**	**3.5**	**1.68**	1.27	0.96	**2.4**	**0.98**	**2.57**	**1.23**

Note: Bold indicates significance at $\alpha < 0.05$.

3.17 points from pre-Institute to post-Institute and 1.74 points from pre-Institute to the one-year follow-up. Using Cohen's d as a measure of practical significance, two of the five features (*learner formulates explanations from evidence* and *learner communicates and justifies explanations*) had medium effect sizes from pre to post administrations (.74 and .59, respectively) while the inquiry feature, *learner connects explanations to scientific knowledge*, showed a large effect size (1.1) pre to post and a medium effect size pre to follow-up (.56). Teachers' self-efficacy for the inquiry feature, *learner gives priority to evidence in responding to questions*, did not change significantly over the programme.

Outcome expectancy

The mean differences and standard deviations for the outcome expectancy scale across all essential features of inquiry are presented in Table 5. Additionally, the multilevel growth model results are presented in Table 6, with all **bolded** estimates indicating statistical significance at the 0.05 level. In this table, the intercept shows the overall mean for each of the five areas of inquiry for the pre-Institute controlling for gender, years of teaching, free/reduced lunch, and school support. For teacher outcome expectancy, two of the five essential features of inquiry were significantly impacted by the teachers' perceived level of school support (*learner engages in scientifically oriented questions and learner formulated explanations from evidence*), in comparison to four out of five inquiry features for self-efficacy.

The teachers' outcome expectancy only significantly changed across two essential features of inquiry (*learner connects explanations to scientific knowledge* and *learner communicated and justifies explanations*) pre to post, with the average growth between pre to post being 1.13 and 2.52 points, respectively (Table 6). These effects were not sustained over time, as the estimates for Time 3 were not significant. Upon close inspection of the means and Figure 2, it seems that the means for three of the five areas (*learner engages in scientifically oriented questions*, *learner formulates explanations from evidence*, and *learner communicates and justifies explanations*) were very close to the maximum value,

Table 5. Mean values and mean differences across essential areas of inquiry for outcome expectancy.

Inquiry feature	Mean (SD) pre	Mean (SD) post	Mean (SD) follow-up	Mean difference post-pre (SD)	Mean difference follow-post (SD)	Mean difference follow-pre (SD)
Learner engages in scientifically oriented questions	31.64 (4.65)	30.92 (6.42)	31.16 (5.51)	−0.72 (6.09)	0.24 (6.04)	−0.48 (4.72)
Learner gives priority to evidence in responding to questions	31.32 (4.63)	31.12 (5.12)	30.84 (5.04)	−0.20 (4.81)	−0.28 (3.92)	−0.48 (3.68)
Learner formulates explanations from evidence	28.44 (3.66)	29.24 (4.20)	29.32 (3.52)	0.80 (4.72)	0.08 (4.18)	0.88 (2.98)
Learner connects explanations to scientific knowledge	16.44 (1.89)	17.56 (2.35)	16.96 (1.81)	1.12 (2.01)	−0.60 (1.98)	0.52 (1.50)
Learner communicates and justifies explanations	31.44 (4.44)	33.04 (5.78)	32.00 (4.79)	1.60 (6.22)	−1.04 (6.15)	0.56 (4.23)

Table 6. Multilevel modelling results for outcome expectancy in each area of scientific inquiry.

	Learner engages in scientifically oriented questions		Learner gives priority to evidence in responding to questions		Learner formulates explanations from evidence		Learner connects explanations to scientific knowledge		Learner communicates and justifies explanations	
	Coeff.	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*
Intercept	**25.3**	**6.75**	**26.44**	**6.57**	**23.79**	**4.35**	**12.87**	**2.62**	**25.36**	**5.91**
Time 2	0.04	1.05	0.22	0.82	1.04	0.82	**1.13**	**0.39**	**2.52**	**1.04**
Time 3	−0.7	1.05	−0.22	0.82	1.00	0.82	0.52	0.39	0.87	1.04
Gender	−2.16	2.87	−1.85	2.79	−2.34	1.84	−0.78	1.11	−3.32	2.51
Years teaching	−0.14	0.11	−0.13	0.11	−0.06	0.07	−0.02	0.04	−0.1	0.1
FRL	−0.05	0.04	−0.03	0.04	−0.04	0.03	−0.01	0.02	−0.03	0.04
School support	**3.87**	**2.01**	2.91	1.96	**2.76**	**1.3**	1.52	0.78	3.25	1.76

Note: Bold indicates significance at $\alpha < 0.05$.

and thus a ceiling effect could have impacted these results (Hessling, Traxel, & Schmidt, 2004; Kraska, 2010). However, teachers felt least confident about changing student outcomes in relation to the essential inquiry feature, *learner connects explanations to scientific knowledge,* with this feature having the lowest mean values over the three TSI administrations (Figure 2). Using Cohen's d, this feature was also the only one to show a medium effect size (.53) and only from pre to post administration. All other essential feature effect sizes for outcome expectancy were small (.28 or below).

Unit implementation

Teachers were asked on the end-of-year survey to indicate how much of the two-week inquiry units taught during the Institute they implemented in their classroom. One

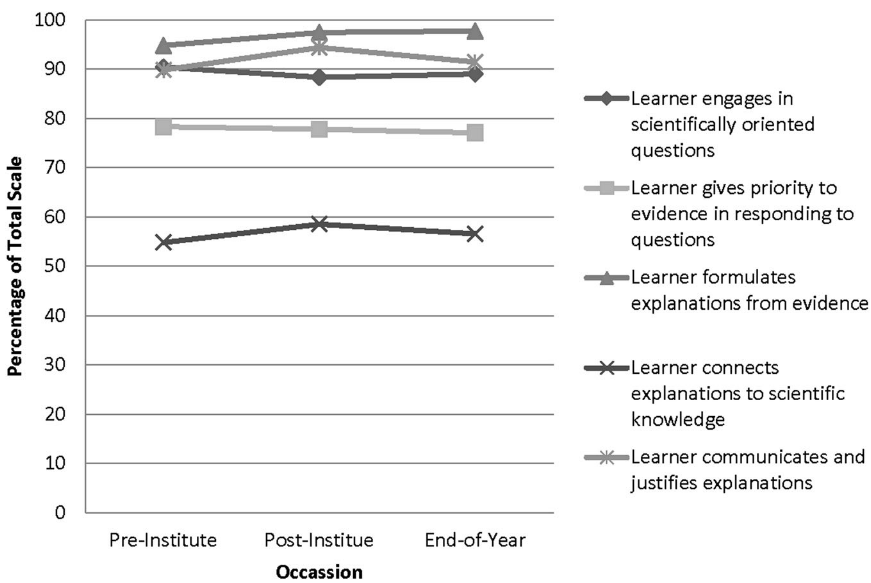

Figure 2. Trajectory of mean percentages for outcome expectancy across time for each essential area of inquiry.

teacher (Marie[1]) indicated that she did not implement any of the unit. Of the remaining 24 teachers, 25% indicated teaching 1–3 lessons, 42% indicated teaching 4–7 lessons, and 33% indicated teaching more than 7 lessons. Sixty-eight per cent of the teachers ($n = 17$) described that they perceived a positive difference in their students' learning during this unit as compared to their typical instruction in the end-of-year survey. Teachers explained this difference in many ways such as through these representative comments: 'They [students] wanted to learn more'; 'They constantly apply their knowledge to new situations at home and tell me about them'; 'I could see the "aha" moments', and 'I could see the results on assessments'. Teachers also described their students as more engaged and as more active participants in observing and explaining the science content.

Seven teachers stated that they were 'uncertain' about whether their students' learning was different. For example, Sue stated 'With students working at different levels it was somewhat challenging.' Christina wrote, 'I think that by asking them to think more, they ended up creating more misconceptions so it took longer to break down the incorrect thinking.' These representative comments show that some teachers were still not connecting inquiry teaching with student learning.

EQUIP

EQUIP factors

The teachers' mean EQUIP scores for each of the five factors increased from pre-institute lesson to post-institute lesson. In order to understand the change from pre to post on each EQUIP factor, regression analyses controlling for demographic variables was used for each of the factors (Instruction, Discourse, Assessment, Curriculum, and Total Inquiry Level). A Mixed Linear Model was conducted on each of the factors (Instruction, Discourse, Assessment, Curriculum, and Total Inquiry Level) to determine if scores changed significantly from pre- to post-test controlling for teacher gender, experience, free/reduced lunch, and perceived school support. The mean, standard deviations, and the effect size (Cohen's d) for each of the subscales are presented in Table 7. The model estimates and significance are presented in Table 8 for each of the subscales. Out of the five subscales, two of the subscale changes from pre-Institute to post-Institute were significant (Discourse and the Total Level). The covariates (teacher gender, experience, free/reduced lunch, and perceived school support) did not significantly impact the results with the exception of free/reduced lunch significantly impacting the Assessment score.

Comparison of TSI and EQUIP

We were also interested in understanding how teachers' different levels of self-efficacy and outcome expectancy compared to their quality of inquiry instruction or EQUIP scores over

Table 7. Descriptive statistics and practical significance for EQUIP subscales.

Variables	Mean (SD) pre	Mean (SD) post	Practical significance Cohen's D
Instructional	2.50 (0.51)	2.74 (0.40)	0.53
Discourse	2.03 (0.55)	2.41 (0.55)	0.69
Assessment	2.05 (0.50)	2.26 (0.44)	0.45
Curriculum	2.31 (0.66)	2.43 (0.45)	0..07
Total	2.22 (0.50)	2.46 (0.37)	0.55

Table 8. Mixed linear model results for pre- and post-EQUIP scores and subscores.

	Equip total		Instruction		Discourse		Assessment		Curriculum	
	Coeff.	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*	*Coeff*	*SE*
Intercept	**49.53**	**9.35**	**14.01**	**2.63**	**11.57**	**3.07**	**11.75**	**2.61**	**12.19**	**2.32**
Time 2	**5.52**	**2.46**	1.26	0.69	**2.13**	**0.81**	1.26	0.69	0.87	0.61
Gender	−2.16	3.96	−1.38	1.11	0.65	1.30	−0.05	1.10	−1.39	0.98
Years teaching	0.04	0.16	<0.01	0.04	0.02	0.05	0.03	0.04	−0.02	0.04
FRL	−0.12	0.06	−0.03	0.02	−0.02	0.02	**−0.03**	**0.02**	−0.04	0.01
School support	−0.33	2.78	0.10	0.78	−0.30	0.91	−0.02	0.78	−0.11	0.69

Note: Bold indicates significance at $\alpha < 0.05$.

time. In order to look at this relationship, the participants were split into either high- or low-efficacy groups; one for outcome expectancy and one for self-efficacy. High and low splits were determined using the median of the distributions from the total scores for both scales on the pre-Institute TSI. Since the sample size was small ($n = 12$ and 13) in each group, a visual analysis was used to examine these effects (Figure 3).

In Figure 3, the graph of the teachers' EQUIP scores from pre-Institute to post-Institute by initial self-efficacy and outcome expectancy levels shows that the teachers with higher self-efficacy and outcome expectancy performed better on the EQUIP before the Institute than teachers with initially lower self-efficacy and outcome expectancy scores. After the programme, teachers initially low in both self-efficacy and outcome expectancy (the trends are similar for both types of efficacy) were comparable to the teachers who were initially high in efficacy. Additionally, both of these groups performed better than they did before the summer institute. Looking at demographic differences in Table 9 between these two groups of teachers, we find that the teachers with initial low self-efficacy had more years of teaching experience ($M = 16.0$, $SD = 10.0$, $n = 13$) when compared to teachers in the high initial self-efficacy group ($M = 7.3$, $SD = 4.2$, $n = 12$). We investigated teachers' perceived mean School Support scores and found that initially low-efficacy teachers also had lower mean (with standard deviation in parentheses) School Support scores

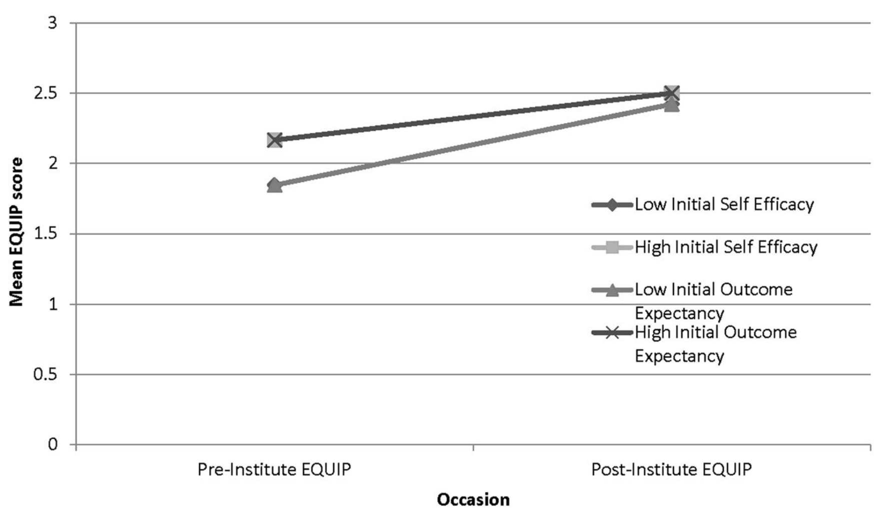

Figure 3. Trajectory of means for EQUIP scores across time by teachers' initial self-efficacy and outcome expectancy level.

Table 9. Teachers' perceived school support scores in relation to their initial self-efficacy and outcome expectancy scores and years of teaching experience.

Teacher pseudonym	Initial SE and OE level	Grade level	Years of teaching experience	Average school support pre-institute	Average school support post-institute
Betty	Low SE and OE	6	26	2.63	2.75
Caitlin	Low SE and OE	7	29	2.88	3.38
Diane	Low SE and OE	8	3	2.88	3.00
George	Low SE and OE	8	4	3.63	3.25
Julie	Low SE and OE	7	7	3.13	2.75
Lori	Low SE and OE	7	3.5	2.75	2.63
Lindsey	Low SE and OE	7	26	3.63	3.88
Leslie	Low SE and OE	6	22	2.88	3.38
Marie	Low SE and OE	8	25	2.00	2.25
Pam	Low SE and OE	6	14	1.75	2.13
Melissa	Low SE and OE	8	27	–	3.50
Average			16.95	2.81	2.99
Laura	Low SE/High OE	7	11	2.38	3.00
Sue	Low SE/High OE	6	11	2.38	2.63
Mary	High SE/Low OE	7	7	2.88	3.13
Patty	High SE/Low OE	8	15	2.63	2.63
Average			11.00	2.56	2.85
Mark	High SE and OE	8	7	2.88	3.00
Courtney	High SE and OE	8	2	2.88	3.00
Carrie	High SE and OE	6	8	3.38	3.88
Connie	High SE and OE	8	11	3.50	4.00
Ed	High SE and OE	7	4	2.63	2.88
Jaime	High SE and OE	7	13	3.13	3.13
Tina	High SE and OE	8	8	3.88	3.50
Angie	High SE and OE	6	5	3.13	3.00
Jessica	High SE and OE	6	1	2.75	3.38
Christina	High SE and OE	8	7	2.75	2.50
Average			6.60	3.09	3.23
Overall average		11.86	2.89	3.06	

than teachers with higher initial self-efficacy ($M = 2.7$ [0.57], 3.0 [0.38] respectively, Table 9). At the end-of-year questionnaire administration, teachers' perceived mean School Support score had significantly increased from the pre-institute administration ($t = 2.43$, $df = 23$, $p = .023$, CI [.286, .023]). This increased feeling of school support might be an indication of a broader positive effect from participating in the PD (greater agency, less administration pressure to improve instruction) or an indication of other positive school-level changes that were not part of this study (Riveros, Newton, & Burgess, 2012). Thus, this conjecture needs further research support.

Essential inquiry features

As described above, teachers increased their personal efficacy for four of the five essential inquiry features after the Institute. We investigated how the teachers described changes in their instruction through the questionnaires after the Institute (post) and at the EOY in regard to their use of questions and evidence-based explanations. We coded teachers by initial efficacy level (high, mixed, low) but did not find any significant patterns or differences when looking across categories. Four teachers were identified as mixed, having one high-efficacy area and one low-efficacy area. For example, Laura was in the low initial self-efficacy group, but had high initial outcome expectancy. Teacher quotes are presented with a high (H), mixed (M), or low (L) initial efficacy designation for reference, but no further divisions were made in the data analysis.

Questioning

Participating teachers' responses to questionnaire items related to instructional growth or changes to instruction focused on their improved questioning skills. Specifically, teachers' responses focused on how they had learnt to facilitate inquiry experiences through teacher-guided questions that helped their students to learn science content. For example, Julie stated, 'I grew the most in giving time for the students to explore and observe. I also grew in asking better questions' (post, L). Jessica said, 'I am more comfortable asking inquiry-based questions requiring my students to think and apply what they know' (post, H). These sentiments continued through the academic year and were commonly stated in the end-of-year questionnaire. For example, George stated, 'I do more questioning and wait for my students to answer and provide the material instead of having them just copy down the notes I have written … they are encouraged to apply the material to new problems' (EOY, L).

Teachers also focused on asking students questions that required their students to think at higher cognitive levels. Jaime stated,

> I have gained the ability to ask more open-ended questions of the students. Inquiry is a major part of learning, and I have not been focused on allowing the students to come up with their own answers, right or wrong. (post, H)

Jaime continued with this thinking at the end of the year when she stated: 'I use more open-ended questions. The question of "why?" and "How does this make sense?" etc. are ways that I am gathering information from my students' (EOY, H). Carrie described similar changes to her instruction at the end of the year: 'I used more open-ended questioning and made students explain their answers. They have to back up their answer with information that they have learned … hands-on yes, but also actively engage in thinking while completing the activity' (EOY, H).

Teachers also described how they would stop using some ineffective questioning strategies such as using: 'fill in the blank oral questioning techniques' (Caitlin, post, L) or 'guess the word I am thinking about' (Julie, post, L) that were often the focus of reflections after the practice-teaching sessions in which teachers were observed using these strategies. For example, Lindsey stated, 'the reflection sessions were very helpful … We were reminded about wait time and our questioning techniques' (post, L).

Questioning was also noted as a challenge for many of the teachers. Lindsey stated that her greatest inquiry teaching challenge was, 'remembering to ask why and keep asking questions. Not doing for the students; I believe I help them too much and I need to lead them into learning the content' (post, L). Laura similarly described her challenge as, 'Not just answering students' questions, but asking questions to help lead them to their own answer' (post, M). Therefore, although teachers described their questioning skills as improving, they were not yet confident that they would not 'revert back' (Caitlin, post, L) to their old more teacher-centred instructional strategies.

When describing their questioning skills, teachers focused on moving from more teacher-centred instruction to student-centred instruction. For example, Courtney stated that her greatest growth during the institute was 'Planning lessons where the students are able to come up with the questions and trying to figure it out on their own with limited help from the teacher' (post, H). Similarly, Jessica described how she would focus on 'less giving of the answers and more exploration on the students' end' (post, H). At the end of the year, Connie stated how she now has her 'students come up with answers rather than me, students have a larger role in the class discussions' (EOY, H). As recorded with the EQUIP Discourse subscale, teachers were questioning at the Developing Inquiry level with a focus on creating inquiry experiences in which students could explore content with teacher guidance.

Each content curriculum unit had an overarching question that teachers introduced to students and came back to investigate throughout the unit (e.g. How can I build a house that will keep me cool in our state?); these questions were not labelled as scientific questions, but rather driving or essential questions. Teachers' self-efficacy towards the essential inquiry feature, 'learner gives priority to evidence in responding to questions' (NRC, 2000, p. 29), did not significantly change during the programme and was only indirectly mentioned through teachers' descriptions of having students provide evidence-based explanations; however, teachers did not explicitly link evidence and questioning in their survey responses. Their discussion of improvements in their questioning skills remained general and connected to establishing a more student-centred environment in their classrooms.

Evidence-based explanations

As described previously, PD instructors modelled the use of two structured protocols for helping students to develop evidence-based explanations: POE (White & Gunstone, 1992) and Claim, Evidence, and Reasoning (CER, McNeil & Krajcik, 2008). Teachers tried these strategies during the practice-teaching and took them back to their classrooms. When asked in the post-institute questionnaire what new instructional strategies they would take back to their classroom from the institute, these two strategies were mentioned most (POE was mentioned 13 times, CER eight times) along with cooperative learning strategies (10 times). George stated:

> I plan on using more of a POE approach while doing labs. I want to limit the procedures I give students and use more guiding questions to lead them into finding the correct results using methods they created. I do plan to use more graphing of data, looking at the class results, analyzing the data, and have them understand that graphs tell a picture. I did not do this very much in the past. (post, L)

Similarly, Diane said, 'I will use the POE model in developing ALL of my lessons next year. Students can come in and make predictions during their "get started" activity; make observations on data, and then explain through pictures/writing/presentations' (post, L). Lindsey described how the practice-teaching allowed her to gain confidence with the using the CER method:

> Through the practice we were able to get in this institute I understand inquiry to be student-oriented activities where the students are held responsible for their learning. It is about the students making a claim as to what they believe and supported it with evidence and then developing an understanding of the concept that was being taught. (post, L)

Even at the EOY, teachers continued to emphasise their use of these protocols. For example, Diane stated: 'Predicting, Observing, and Explaining (POE), I think these are key components to any inquiry lesson. I think the key is having students come up with their OWN explanations for science concepts' (EOY, L). Christina stated, 'I have had

Table 10. Teachers' descriptions of the characteristics of effective inquiry instruction before and after the summer institute and at the EOY.

Effective inquiry lesson characteristics	Pre-institute total	Post-institute total	End of year total
Student-centred, teacher-guided	8	11	5
Data analysis	6	11	3
Investigation/problem-solving	6	0	4
Engaging for students	5	2	0
Student relevant content	3	6	3
Hands on, minds on	2	0	1
Student discovery	2	6	4
Student collaboration/discussion	2	5	2
Students' questioning	2	11	5
Teacher open-ended questions (not one right answer)	2	4	4
Student learning/understanding of content	2	10	6
Evidence-based explanations (includes CER)	1	12	7
Connect to students prior knowledge/misconceptions	1	4	6
Phenomena first activity/laboratory	0	2	3
POE	0	7	4

Note: Numbers equal the number of times the characteristics were written by different teachers, not the number of teachers, teachers often mentioned several features within their inquiry definition.

them make more predictions before we begin units about the concepts that we are studying. I have used the P-O-E method quite a bit to do this' (EOY, H).

After the Institute, teachers' inquiry definitions changed to include the importance of evidence-based explanations, a greater focus on student learning, students' questioning, and POE. Table 10 summarises teacher's definitions of inquiry teaching before and after the summer Institute and at the end of the year. Melissa stated,

> Prior knowledge does help, but is not always necessary. Being able to look at data and come up with logical explanations is an important tool. But, looking at the data and not coming up with the correct response is also a great way to teach something. (EOY, L)

Similarly, Tina stated:

> An inquiry lesson … is student-centered by having them predict why it's happening. It is driven by students deciding how they will test their prediction and gain data/observations. And concluded with students using all of the above to decide on an explanation/ reasoning based on real data and science and revisiting their prediction. (post, H)

Thus, after attending the Institute, teachers more often connected inquiry with evidence-based explanations and implemented the POE and CER protocols with their students.

Discussion

The participating teachers' significantly increased their self-efficacy for designing inquiry lessons in which *students engage in scientifically oriented questions*, *formulate explanations from evidence*, *connect explanations to scientific knowledge*, and *communicate and justify their explanations* after the Institute. Through the inquiry PD model, we emphasised the importance of students learning scientific argumentation and provided the teachers with two structured protocols to improve students' scientific explanations: Claim, Evidence, and Reasoning (McNeil & Krajcik, 2008) and POE (White & Gunstone, 1992). The use and reinforcement of these two structured protocols during the pedagogy, content, and practice-teaching sessions may have made these complex inquiry tasks more accessible for teachers. Therefore, if self-efficacy involves an analysis of the task and the teaching context (Tschannen-Moran et al., 1998), the scaffolding of these tasks through the use of the structured protocols and positive reinforcement through practice and reflection may have helped the teachers to gain confidence in their capabilities to perform these specific inquiry features. Similarly, Palmer (2011) reported increases in teachers' personal self-efficacy for teaching science after participation in a PD programme that had teachers practise a specific investigative sequence. Like our programme, the teachers in Palmer's study experienced this protocol as a 'student' during the PD programme and as a teacher in their own classroom.

Overall, teachers in our programme did not increase their self-efficacy for the essential feature *learner gives priority to evidence in responding to questions*. Teachers often associate inquiry with teacher and student questions (Lotter, Harwood, & Bonner, 2007; Roehrig & Luft, 2004; Wallace & Kang, 2004) and the teachers in our programme may have believed that they were already capable of using this inquiry skill before the programme. However, after participating in the Institute, teachers wrote in their final reflections and questionnaires about how they improved their questioning skills through the use of more open-ended higher level questions. The teachers might have initially overestimated their efficacy in regard to their questioning skills and thus did not show improvements after the Institute. Tschannen-Moran et al. (1998) describe this as a 'reality shock' as teachers new to teaching or to an instructional strategy 'confront the complexity of the teaching task' (p. 232). Teachers also described in the end-of-year questionnaires how their questioning skills were still developing as they transitioned from a teacher-directed classroom culture to a more student-centred culture in which students were given time to explore, collect, and analyse data to develop evidence-based science explanations. Our observations of teachers classrooms and EQUIP discourse scores within the Developing Inquiry range (Level 2) also support the notion that our teachers' questioning skills were improving, but not yet exemplary (Level 3).

The teachers' outcome expectancy for TSI increased over the PD programme, but only slightly. The only significant changes in the teachers' outcome expectancy were for the *learner connects explanations to scientific knowledge* and the *learner communicates and justifies explanations* and only from pre-Institute to post-Institute. Teachers may have increased their outcome expectancy for these essential features due to an increase in their own content knowledge after the Institute or from observing students making content connections during the practice-teaching sessions. As the teachers returned to their classrooms and the contextual constraints of their schools (pacing guides, high-stakes tests, unsupportive colleagues, etc.), their efficacy for teaching inquiry may have been lowered. Similarly, Roehrig et al. (2007) found that chemistry teachers' implementation of inquiry curriculum was influenced by several school-specific issues including school-level administrator support, class schedules, and other ongoing reform initiatives.

Alternatively, the teachers in our programme may have overestimated their efficacy to teach and help students learn through inquiry (ceiling effect) before the Institute: not having a complete understanding of the instructional strategy (Lotter, Harwood, & Bonner, 2007; Settlage et al., 2009) or an accurate conceptual framework for characterising good inquiry teaching (White & Frederiksen, 2000). Misconceptions about inquiry are common in the literature with teachers equating inquiry with questioning or hands-on learning and not with data collection or scientific argumentation (Lotter, Harwood, &

Bonner, 2007; Wallace & Kang, 2004). A common statement by our teachers on the end-of-year questionnaire was that they gained a true understanding of inquiry teaching. Our data show an increase in the complexity of teachers' inquiry definitions after the Institute with an increased focus on evidence-based explanations, students' prior knowledge, data analysis, and student questioning. This finding supports the conclusions of Capps and Crawford (2013b) that to enact inquiry instruction, teachers need to reach a certain level of understanding about this complex instructional strategy.

The complexity of inquiry teaching extends beyond teachers' conceptual understanding of inquiry and includes the difficulties associated with teaching students how to think differently, participate in productive classroom discourse, and replace their prior misconceptions with new science knowledge. Windschitl (2002b) described these complexities as four dilemmas (conceptual, pedagogical, cultural, and political) that teachers need to overcome to be able to teach within a constructivist framework. During our Institute, through practice-teaching and observing inquiry models, our teachers changed or re-aligned their conceptions of inquiry to be more in line with the visions of inquiry within reform documents (NRC, 2000). When they returned to their classrooms and tried 'truer' inquiry-based practices, some teachers described that the inquiry strategies revealed student misconceptions. This may have led them to believe that they had less ability to have a positive impact (outcome expectancy) on student learning because they did not yet have the skills needed to change these misconceptions given their still developing level of inquiry teaching (Blanchard et al., 2010).

Researchers have argued that changes in outcome expectancy lag behind changes in self-efficacy (Lakshmanan et al.'s, 2011). Guskey (1986) supported this notion when he described how teachers might need to observe substantial changes in student learning before changes in teacher beliefs and practice are solidified. In spite of emerging evidence, our study findings are limited by the analysis of only one inquiry lesson before and after PD participation. Longitudinal studies are needed to confirm these assumptions. Furthermore, other commonly reported teacher beliefs, such as teachers believing that students are incapable of learning through inquiry or teachers feeling time constraints to teach students using inquiry (Lotter, Harwood, & Bonner, 2007; Blanchard et al., 2009), may be responsible for the lack of change in our teachers' outcome expectancy values.

Our PD model that engages teachers in practice-teaching and reflection is supported by related research (Ingvarson et al., 2005) which highlights that PD programme teacher efficacy outcomes are most influenced by teachers' beliefs that a programme will positively influence their students' learning. As Ingvarson et al. (2005) state, 'Programs that model effective practice and invite teachers to try them out tend to be more successful than programs that devote resources primarily to changing attitudes first' (p. 16). Similarly, Capps et al. (2012), in their review of inquiry PD programmes, described a need for programmes that 'provide a structure for challenging teachers to examine their knowledge and beliefs and reflect on their teaching practice, allow teachers the opportunity to experience authentic scientific inquiry in meaningful contexts similar to how they will teach in their classrooms' (p. 307). Thus, our programme provides additional support for PD models that provide opportunities for teachers to practise new instructional strategies in supportive reflective contexts.

We did find an association between teacher self-efficacy (4 of 5 features), teacher outcome expectancy (2 of 5 features), and teacher's perceived level of school support.

This is similar to other research that reveals that school context and school-level support are factors impacting general teacher self-efficacy (Ashton, 1984; Hoy & Spero, 2005; Knoblauch & Hoy, 2008; Tschannen-Moran & Hoy, 2007) as well as teachers' use of inquiry-based instruction (Roehrig et al., 2007). Schriver and Czerniak (1999) provide a good example of the impact that school support can have on teaching efficacy. Middle school teachers in their study who had strong knowledge of middle-level curriculum and taught in schools with a team teaching and planning structure had higher outcome expectancy (STEBI; Riggs & Enochs, 1990) than teachers who taught in junior high school settings that placed teachers in more isolated teaching environments. Programmes, like ours, that increase the amount of teacher collaboration during a PD experience have also been shown to increase teacher efficacy (Henson, 2001; Lumpe, Czerniak, Haney, & Beltyukova 2012).

Overall, our PD improved the inquiry instruction of participating teachers regardless of their initial self-efficacy level. Settlage et al. (2009) highlight self-doubt as a powerful force that pushes teachers to perceive a need to change and reflect further on their instruction. Likewise, Lakshmanan et al. (2011) found that teachers with lower initial self-efficacy and outcome expectancy maintained a lower level of inquiry instruction in their classroom. Our results, although over only a one-year period, show teachers' inquiry skills converging (our post EQUIP scores for both low- and high self-efficacy and outcome expectancy scores were 2.4–2.5) rather than diverging based on teachers' initial efficacy beliefs. Our use of practice-teaching during the Institute may have forced teachers who would have normally not tried inquiry teaching to experiment and succeed in a low-risk context (summer enrichment programme) with this teaching strategy. Other studies suggest that additional changes in instruction (e.g. moving to Level 3 inquiry on the EQUIP) may require multiple year interventions or additional school-based support due to the complexity of inquiry-based instruction (Luft, 2001; Marshall, 2013; Roehrig et al., 2007).

Conclusion

Our PD programme provided teachers with an opportunity to learn science content and inquiry-based teaching skills in a controlled and supportive environment. The teachers planned, practice-taught, and reflected on this teaching with peers and pedagogy experts who were all working towards the same goal, improved inquiry instruction. This collective goal and the supportive environment may have created a form of collective teacher efficacy during the summer programme (Goddard et al., 2000). We saw gains in the teachers' self-efficacy and outcome expectancy towards teaching inquiry immediately after the Institute that diminished (for all but one feature) once the teachers returned to their school contexts. Tschannen-Moran et al. (1998) describe a 'cohesive [school] culture' that increases collective efficacy as '… one that is orderly, with a strong press for academic achievement, where administrators are responsive to teachers' concerns and encourage them to try new ideas, and where teachers encourage one another in their attempts to address student needs' (p. 222). Collective teacher efficacy within schools has been shown to be predictive of student achievement (Goddard et al., 2000). Creating this supportive environment both within a PD programme but more importantly within teachers' schools might be essential to long-term changes in teachers' instruction and student learning. The significant association we found between our teachers' increased efficacy and perceived level of school support also reinforces this conjecture. Thus, our PD model, however successful in moving teachers towards inquiry during the summer Institute, might be improved through the recruitment of school-based teacher teams or through a focus on whole-school PD to increase support and change during the academic year.

Providing time for teachers to practice-teach inquiry skills with students and gain reflective feedback on how to improve their instruction during a summer PD was valued by the participating teachers and cited by some as reasons for their increased confidence with using these skills. As stated by Palmer (2011):

> Unfortunately, in situ modeling is probably not a viable activity in the daily life of teachers, as it requires a commitment by experts, who may not necessarily be readily available. However, in the context of professional development programs, in situ feedback is a technique that has great potential to enhance teacher self-efficacy. (p. 594)

Our PD programme provides additional evidence to support this conclusion in regard to inquiry science teaching. Our study also provided some evidence that teachers with initially low inquiry teaching efficacy, if provided time to practise inquiry skills within a safe PD environment, can increase the quality of their inquiry teaching as much as teachers with initially high self-efficacy. Further evidence will need to be collected on a larger group of teachers over a longer time period to further substantiate these findings.

Note

1. All teachers' names are pseudonyms.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the US Department of Education Mathematics and Science Partnerships [Grant number H63010008209].

References

Akerson, V., Cullen, T., & Hanson, D. (2009). Fostering a community of practice through a professional development program to improve elementary teachers' views of nature of science and teaching practice. *Journal of Research in Science Teaching*, *46*, 1090–1113. doi:10.1002/tea.20303

Amanti, K. (2000). Constructing scientific models in middle school. In J. Minstrell & E. H. van Zee (Eds.), *Inquiry into inquiry learning and teaching in science* (pp. 316–330). Washington, DC: American Association for the Advancement of Science.

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. 807–830). Mahwah, NJ: Lawrence Erlbaum Associates.

Ashton, P. T. (1984). Teacher efficacy: A motivational paradigm for effective teacher education. *Journal of Teacher Education*, *35*(5), 28–32. doi:10.1177/002248718403500507

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W.H. Freeman and Company.

Blanchard, M. R., Southerland, S. A., & Granger, E. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education*, *93*, 322–360. doi:10.1002/sce.20298

Blanchard, M. R., Southerland, S. A., Osborne, J., Sampson, V., Annetta, L., & Granger, E. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, *94*, 577–616. doi:10.1002/sce.20390

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research for education* (3rd ed.). Boston: Allyn and Bacon.

Brown, J. S., Collins, A., & DuGuid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42. doi:10.3102/0013189X018001032

Capps, D. K., & Crawford, B. A. (2013a). Inquiry-based instruction and teaching about nature of science: Are they happening? *Journal of Science Teacher Education*, *24*, 497–526.

Capps, D. K., & Crawford, B. A. (2013b). Inquiry-based professional development: What does it take to support teachers in learning about inquiry and nature of science? *International Journal of Science Education*, *35*, 1947–1978.

Capps, D. K., Crawford, B. A., & Constas, M. A. (2012). A review of empirical literature on inquiry professional development: Alignment with best practices and a critique of the findings. *Journal of Science Teacher Education*, *23*, 291–318. doi:10.1007/s10972-012-9275-2

Dira-Smolleck, L. A. (2004). *The development and validation of an instrument to measure preservice teachers' self-efficacy in regards to the teaching of science as inquiry* (Unpublished doctoral dissertation). The Pennsylvania State University.

Enochs, L. G., Scharmann, L. C., & Riggs, I. M. (1995). The relationship of pupil control to preservice elementary science teacher self-efficacy and outcome expectancy. *Science Education*, *79*, 63–75. doi:10.1002/sce.3730790105

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

Feldman, A. (2000). Decision making in the practical domain: A model of practical conceptual change. *Science Education*, *84*, 606–623. doi:10.1002/1098-237X(200009)84:5<606::AID-SCE4>3.0.CO;2-R

Fishman, B., Marx, R., Best, S., & Revital, T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, *19*, 643–658.

Goddard, R., Hoy, W., & Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, *37*, 479–507. doi:10.3102/00028312037002479

Greeno, J. G. (1997). On claims that answer the wrong questions. *Educational Researcher*, *26*(1), 5–17.

Gregoire, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, *15*, 147–179. doi:10.1023/A:1023477131081

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher*, *15*, 5–12. doi:10.3102/0013189X015005005

Henson, R. A. (2001). The effects of participation in teacher research on teacher efficacy. *Teaching and Teacher Education*, *17*, 819–836. doi:10.1016/S0742-051X(01)00033-6

Hessling, R., Traxel, N., & Schmidt, T. (2004). Ceiling effect. In M. Lewis-Beck, A. Bryman, & T. Liao (Eds.), *Encyclopedia of social science research methods* (pp. 106–107). Thousand Oaks, CA: SAGE.

Horizon Research, Inc. (2006). *Local systemic change through teacher enhancement questionnaire*. Retrieved from http://www.horizon-research.com/local-systemic-change-through-teacher-enhancement-science-6-12-teacher-questionnaire/

Hoy, A. W., & Spero, R. B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education*, *21*, 343–356. doi:10.1016/j.tate.2005.01.007

Ingvarson, L., Meiers, M. & Beavis, A. (2005). Factors affecting the impact of professional development programs on teachers' knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives*, *13*(10). Retrieved July 28, 2016, from http://epaa.asu.edu/epaa/v13n10/

Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2010). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*, 21–43.

Kleine, K., Brown, B., Harte, B., Hilson, A., Malone, D., & Moller, K. (2002). Examining inquiry. *Principal Leadership*, *3*(3), 36–39.

Knoblauch, D., & Woolfolk Hoy, A. (2008). 'Maybe I can teach those kids.' The influence of contextual factors on student teachers' efficacy beliefs. *Teaching and Teacher Education*, *24*, 166–179. doi:10.1016/j.tate.2007.05.005

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall.

Krajcik, J. S., Blumenfeld, P., Marx, R. W., & Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Mistrell, & e. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 283–315). Washington, DC: American Association for the Advancement of Science.

Kraska, M. (2010). Repeated measures design. In N. Salkind (Eds.), *Encyclopedia of research design*. (pp. 1244–1248). Thousand Oaks, CA: SAGE.

Lakshmanan, A., Heath, B., Perlmutter, A., & Elder, M. (2011). The impact of science content and professional learning communities on science teaching efficacy and standards-based instruction. *Journal of Research in Science Teaching*, *48*, 534–551. doi:10.1002/tea.20404

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lotter, C., Harwood, H., & Bonner, J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices. *Journal of Research in Science Teaching*, *44*(9): 1318–1347.

Lotter, C., Thompson, S., Dickenson, T., Smiley, W., Bilue, G., & Rea, M. (in press). The impact of a practice-teaching professional development model on teachers' inquiry instruction and inquiryefficacy beliefs. *International Journal of Science and Mathematics Education*, 1–19. doi:10.1007/s10763-016-9779-x

Lotter, C., Yow, J., & Peters, T. (2014). Building a community of practice around inquiry instruction through a professional development program. *International Journal of Science and Mathematics Education*, *12*, 1–23. doi: 07/s10763-012-9391-7.

Loughran, J. (2014). Developing understandings of practice: Science teacher learning. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. *II*, pp. 811–829). New York, NY: Routledge.

Luft, J. A. (2001). Changing inquiry practices and beliefs: The impact of an inquiry-based professional development programme on beginning and experienced secondary science teachers. *International Journal of Science Education*, *23*, 517–534. doi:10.1080/09500690121307

Lumpe, A., Czerniak, C., Haney, J., & Beltyukova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation in professional development and student achievement. *International Journal of Science Education*, *34*, 153–166. doi: 10.1080/09500693.2010.551222.

Magnusson, S., Krajick, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome, & N. G. Lederman (Eds.), *Examining pedagogical content knowledge* (pp. 95–132). Dordrecht: Kluwer Academic.

Marco-Bujosa, L. M., McNeill, K. L., González-Howard, M., & Loper, S. (2016). An exploration of teacher learning from an educative reform-oriented science curriculum: Case studies of teacher curriculum use. *Journal of Research in Science Teaching*, doi:10.1002/tea.21340

Marshall, J. C. (2013, March). *Giving up before the finish line: Teacher transformation resulting in improved student achievement takes time*. Research paper presentation at National Association for Research in Science Teaching Conference, Puerto Rico.

Marshall, J. C., Smart, J., & Horton, R. (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. doi:10.1007/s10763-009-9174-y

McNeil, K., & Krajcik, J. S. (2008). Inquiry and scientific explanations: Helping students use evidence and reasoning. In J. A. Luft, R. L. Bell, & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting*. Arlington: NSTA press.

Metz, K. E. (2000). Young children's inquiry in biology: Building the knowledge bases to empower independent inquiry. In J. Minstrell, & E. van Zee (Eds.), *Inquiring into inquiry in science learning and teaching* (pp. 371–404). Washington, DC: American Association for the Advancement of Science.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: National Academy Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Newman, I., Ridenour, C., Newman, C., & DeMarco, G. M. P. (2003). A typology of research purposes and its relationship to mixed methods. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (pp. 167–188). London: SAGE.

Next Generation Science Standards (NGSS) Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

O'Brien, T. (1992). Science in-service workshops that work for elementary teachers. *School Science and Mathematics*, *92*(8), 422–426. doi:10.1111/j.1949-8594.1992.tb15622.x

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, *66*, 543–578. doi:10.3102/00346543066004543

Palmer, D. (2011). Sources of efficacy information in an inservice program for elementary teachers. *Science Education*, *95*, 577–600. doi:10.1002/sce.20434

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, *29*(1), 4–15.

Raudenbush, S. W. (1993). Hierarchical linear models as generalizations of certain common experimental designs. In L. Edwards (Eds.), *Applied analysis of variance in behavioral science* (pp. 459–496). New York, NY: Marcel Dekker.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J. Sikula (Ed.), *The handbook of research in teacher education* (2nd ed., pp. 102–119). New York, NY: Macmillan.

Riggs, I., & Enochs, L. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*, 625–637. doi:10.1002/sce.3730740605

Riveros, A., Newton, P., & Burgess, D. (2012). A situated account of teacher agency and learning: Critical reflections on professional learning communities. *Canadian Journal of Education*, *35*, 202–216.

Roehrig, G. H., Kruse, R. A., & Kern, A. (2007). Teacher and school characteristics and their influence on curriculum implementation. *Journal of Research in Science Teaching*, *44*(7), 883–907. doi:10.1002/tea.20180

Roehrig, G. H., & Luft, J. A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education*, *26*(1), 3–24. doi:10.1080/0950069022000070261

Ross, J., & Bruce, C. (2007). Professional development effects on teacher efficacy: Results of randomized field trial. *The Journal of Educational Research*, *101*(1), 50–60. doi:10.3200/JOER.101.1.50-60

Schriver, M., & Czerniak, C. M. (1999). A comparison of middle and junior high science teachers' level of efficacy and knowledge of developmentally appropriate curriculum and instruction. *Journal of Science Teacher Education*, *10*, 21–42. doi:10.1023/A:1009472629345

Settlage, J., Southerland, S. A., Smith, L. K., & Ceglie, R. (2009). Constructing a doubt-free teaching self: Self-efficacy, teacher identity, and science instruction within diverse settings. *Journal of Research in Science Teaching*, *46*(1), 102–125. doi:10.1002/tea.20268

Sinclair, B., Naizer, G., & Ledbetter, C. (2011). Observed implementation of a science professional development program for K-8 classrooms. *Journal of Science Teacher Education*, *22*, 579–594. doi:10.1007/s10972-010-9206-z

Singer, J., Lotter, C., Gates, A., & Feller, R. (2011). Exploring a model of situated professional development: Impact on classroom practice. *Journal of Science Teacher Education*, *22*, 203–227.

Smolleck, L. A., & Yoder, E. P. (2008). Further development and validation of the Teaching Science as Inquiry (TSI) instrument. *School Science and Mathematics*, *108*, 291–297. doi:10.1111/j.1949-8594.2008.tb17842.x

Smolleck, L. A., Zembal-Saul, C., & Yoder, E. P. (2006). The development and validation of an instrument to measure preserve teachers' self-efficacy in regard to the teaching of science as inquiry. *Journal of Science Teacher Education*, *17*, 137–163. doi:10.1007/s10972-006-9015-6

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Thompson, C. L., & Zeuli, J. S. (1999). The frame and the tapestry. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession* (pp. 341–375). San Francisco: Jossey-Bass.

Tschannen-Moran, M., Hoy, A., & Hoy, W. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research*, *68*, 202–248. doi:10.3102/00346543068002202

Tschannen-Moran, M., & Hoy, A. W. (2007). The differential antecedents of self-efficacy beliefs of novice and experienced teachers. *Teaching and Teacher Education*, *23*, 944–956. doi:10.1016/j.tate.2006.05.003

Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education*, *25*, 1051–1060. doi:10.1016/j.tate.2009.04.002

Wallace, C. W., & Kang, N. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: An examination of competing belief sets. *Journal of Research in Science Teaching*, *41*, 936–960. doi:10.1002/tea.20032

Weis, A. M. (2013). *2012 National Survey of Science and Mathematics Education: Status of middle school science*. Chapel Hill, NC: Horizon Research.

Wheatley, K. F. (2002). The potential benefits of teacher efficacy doubts for educational reform. *Teaching and Teacher Education*, *18*, 5–22. doi:10.1016/S0742-051X(01)00047-6

White, B., & Frederiksen, J. (2000). Metacognitive facilitation: An approach to making scientific inquiry accessible to all. In J. Minstrell, & E. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 331–370). Washington, DC: American Association for the Advancement of Science.

White, R. T., & Gunstone, R. F. (1992). *Probing understanding*. Great Britain: Falmer Press.

Windschitl, M. (2002a). Inquiry projects in science teacher education: What can investigative experience reveal about teacher thinking and eventual classroom practice? *Journal of Science Teacher Education*, *87*, 112–143.

Windschitl, M. (2002b). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, *72*(2), 131–175. doi:10.3102/00346543072002131

Zwart, R., Wubbels, T., Bolhuis, S., & Bergen, T. (2008). Teacher learning through reciprocal peer coaching: An analysis of activity sequences. *Teaching and Teacher Education*, *24*, 982–1002. doi:10.1016/j.tate.2007.11.003