Assessing Gains in Science Teaching Self-Efficacy After Completing an Inquiry-Based Earth Science Course

Kyle Gray[1,a]

ABSTRACT

Preservice elementary teachers are often required to take an Earth Science content course as part of their teacher education program but typically enter the course with little knowledge of key Earth Science concepts and are uncertain in their ability to teach science. This study investigated whether completing an inquiry-based Earth Science course for preservice elementary teachers improved their self-efficacy beliefs toward science teaching using the Science Teaching Efficacy Belief Instrument (STEBI-B). The results were compared against the number of years the students had been undergraduates. All students improved their science teaching self-efficacy (including the freshmen) but only the juniors and seniors experienced significant gains in outcome expectancy. The observed gains in self-efficacy were likely due to increases in cognitive content mastery—the process of successfully learning new material, and the gain in outcome expectancy were probably influenced by the number of field experiences within their teacher-education program. Geoscience departments should, therefore, offer inquiry-based Earth Science content courses designed specifically for preservice elementary teachers to improve their self-efficacy and content knowledge, thus, meeting the needs of this critical student population. [DOI: 10.5408/14-022.1]

Key words: self-efficacy, preservice elementary teachers, Earth Science content courses, inquiry

INTRODUCTION

Many geoscience departments offer content courses specifically designed for preservice elementary teachers. These students often avoid science courses (Rice and Roychoudhury, 2003), enter our classes with little background in geoscience (Tosun, 2000b), and feel unqualified to teach Earth Science concepts (Plourde, 2002), yet many of them will become the first science teachers our children encounter. The importance of this issue is exacerbated by the increased role of the geosciences (Wysession, 2012) in the Next Generation Science Standards (NGSS) (NGSS Lead States, 2013). Therefore, it is imperative that preservice elementary teachers not only learn geoscience concepts but also feel comfortable with teaching the subject matter.

Self-efficacy is defined as one's confidence toward completing a given task (Bandura, 1977), and during the past 4 decades, educational researchers have applied this construct to various aspects of teacher education. High levels of teacher self-efficacy have been linked to improved student learning (Ross, 1992; Caprara et al., 2006) and teacher job satisfaction (Brouwers and Tomic, 2000; Caprara et al., 2006). Teachers who have a strong sense of self-efficacy also persist longer when working with students who struggle academically (Gibson and Dembo, 1984) and are open to innovative pedagogies (Ghaith and Yaghi, 1997; Tschannen-Moran et al., 1998). In this light, it is imperative that those geoscience educators who work with preservice teachers establish a curriculum that both addresses student-content needs and fosters improved teacher self-efficacy. This study investigated the effect of an inquiry-based, Earth and Space Science content course on science-teaching self-efficacy among preservice elementary teachers.

Received 17 April 2014; revised 18 July 2014 and 22 November 2016; accepted 29 November 2016; published online 14 March 2017.

[1]Department of Earth Science, University of Northern Iowa, 121 Latham Hall, Cedar Falls, Iowa 50614-0335, USA

[a]Author to whom correspondence should be addressed. Electronic mail: kyle.gray@uni.edu. Tel.: 319-273-2809. Fax: 319-273-7124

LITERATURE REVIEW

Self-efficacy is a perception that is not dependent on one's actual ability to achieve the stated task (Bandura, 1984). This means it is possible for someone to believe that they can achieve a given task yet be incapable of doing so. We observe such behavior in the classroom when students believe they earned a high grade on an exam yet actually received a low score. Given that self-efficacy is an internal belief, it can only be measured by asking the participant to evaluate their own ability and cannot be externally observed or measured.

Self-efficacy is also context dependent and domain specific (Pajares, 1996; Klassen et al., 2011). Because judgments of self-efficacy are directly tied to task achievement, researchers must clearly define which actions the participants are being evaluated for. Pajares (1996) noted that a balance must be struck between assessing an ill-defined sense of general self-efficacy and investigating the construct on such a microscopic level that the results lose all practical meaning.

Sources of Self-Efficacy

Bandura (1977) cited four sources that affect a person's self-efficacy beliefs. Two sources center on successful completion of the target task. During *mastery experiences* (called *performance accomplishments* by Bandura), a person actively engages in the target task and experiences success, but during *vicarious experiences*, the person watches a third party successfully complete the target task. The remaining sources focus on factors that motivate a person to successfully complete a given task. V*erbal persuasion* consists of spoken encouragement by a third party, whereas *emotional arousal* occurs when a person's emotions strongly influence his or her ability to compete the given task. Research on the relative strength of these four constructs consistently show that mastery experiences hold the greatest influence on self-efficacy with vicarious experiences, verbal persuasion, and emotional arousal contributing a small or nonsignificant influence on self-efficacy (Bandura, 1997; Mulholland and Wallace, 2001; Britner and Pajares, 2006; van Dinther et al., 2011). It is hypothesized that, after successfully completing a task, students reevaluate their attitudes toward the target subject and adjust their confidence toward that task (Usher and Pajares, 2008). However, positive experiences alone do not guarantee improvements in self-efficacy. Students must cognitively process those experiences in light of personal and environmental factors (Britner and Pajares, 2006), and cultural and social factors may also influence the level to which mastery experiences improve self-efficacy (Usher and Pajares, 2006, 2008).

More recently, Palmer (2006b) argued that preservice teachers tap into additional forms of mastery experiences that in turn influence the student's science teaching self-efficacy. He defined *cognitive content mastery* as "mastery experiences in understanding new content" (p. 339) and *cognitive pedagogical mastery* as learning new ways to teach a given topic. Research on the relative strengths of these additional sources of self-efficacy suggests that both cognitive content mastery and cognitive pedagogical mastery directly improve student self-efficacy toward teaching, with cognitive pedagogical mastery being the most influential (Palmer, 2006b; Bender, 2007; Bautista, 2011; Zundans-Fraser and Lancaster, 2012). These findings suggest that learning new science content can improve teaching self-efficacy, and the findings are consistent with other studies that found a strong relationship between content knowledge and teacher self-efficacy (Muijs and Reynolds, 2002; Cantrell et al., 2003; Palmer et al., 2015). Given that many undergraduate students have had limited Earth Science experiences in their K–12 education (Tosun, 2000b) and are apprehensive toward teaching Earth Science (Plourde, 2002), learning new material in this subject may greatly influence their teaching self-efficacy.

Self-Efficacy and Science Content Courses

Most studies on science-teacher self-efficacy have focused on students enrolled in courses on science teaching methods near the end of their undergraduate program (Watters and Ginns, 1995; Huinker and Madison, 1997; Palmer, 2006b; Bleicher, 2009; Bautista, 2011). In contrast, few studies have investigated the influence of undergraduate courses in science content on science-teaching self-efficacy (Posnanski, 2007; Avery and Meyer, 2012; Palmer et al., 2015). Other studies have found that preservice teachers who complete more undergraduate science courses possess higher levels of science-teaching self-efficacy (Enochs and Riggs, 1990; Ramey-Gassert et al., 1996). Luera and Otto (2005) assessed gains in self-efficacy for students taking three inquiry-based, science-content courses and found that self-efficacy did not significantly increase until students had completed two such courses. However, other studies have not found a link between science-content courses and self-efficacy (Tosun, 2000b; Morrell and Carroll, 2003; Bleicher and Lindgren, 2005). One possible explanation may lie in how those courses were taught. Morrell and Carroll (2003) compared the levels of self-efficacy of students taking a science methods course to students taking lecture-based content courses and found that the students taking the lecture courses did not improve their science-teaching self-efficacy, whereas both Posnanski (2007) and Luera and Otto (2005) found that science content courses built on a constructivist pedagogy did improve student science-teaching self-efficacy.

More recently, several studies have investigated courses that mix lecture and inquiry-driven pedagogies. Palmer et al. (2015) found that students improved their science-teaching self-efficacy after taking a lecture-based course with a weekly tutorial comprising hands-on activities, and Baldwin (2014) found the same for a traditional lecture course with a laboratory section designed specifically for preservice teachers. However, Avery and Meyer (2012) included open-ended laboratory investigations with lectures and found that nearly one-third of their students did not improve their self-efficacy.

Despite this clear connection between science-teaching self-efficacy and the number of undergraduate science courses, caution must be taken when attributing the observed gains to the number of science courses taken. For example, the Morrell and Carroll (2003) data on science-content courses came from freshmen and sophomore preservice teachers, but data from methods assessed juniors and seniors. Similarly, Baldwin (2014) aggregated the students as a single group even though they ranged from freshmen to juniors. As will be discussed below, the number of science courses may correlate with the number of years a student has been at a university. Freshmen are unlikely to have more than one prior course, and seniors have had more opportunities to complete additional science courses. Studies on self-efficacy that include the student university year as a variable are rare. Aydin and Boz (2010) found that seniors hold slightly greater self-efficacy beliefs than do freshmen, sophomores, or juniors, and Bayraktar (2011) found that seniors held significantly greater self-efficacy beliefs than did freshmen.

However, when taken together, all these studies on the link between content courses and self-efficacy found that hands-on, inquiry-based courses were more likely to improve student self-efficacy than were straight lecture and traditional "cookbook" confirmatory laboratories. These studies illustrate the need for science-content courses to address the issue of self-efficacy among preservice elementary teachers, especially in light of those students often entering our classes with low science-teaching self-efficacy, which, in turn, motivates them to avoid teaching science as in-service teachers (Palmer, 2006a)

Preservice Teachers in Geoscience Courses

Within the geosciences, few studies have investigated the relationship between Earth Science content courses and improvements in science-teaching self-efficacy, but the results generally indicated that taking undergraduate geoscience-content courses improved self-efficacy. A few studies (Posnanski, 2007; Slater et al., 2008; Gosselin et al., 2010; Baldwin, 2014) have reported improvements in self-efficacy after students took geoscience courses that used a constructivist pedagogy. Posnanski (2007) also found that self-efficacy did not improve for those students who completed traditional, lecture-based courses. Finally, Stone (2016) analyzed open-ended responses, from 54 students who also participated in the study reported here, and found that students taking an inquiry-based content course attributed their level of self-efficacy to their content knowledge. Furthermore, these students cited an increase in content knowledge as a major factor in their improved self-efficacy beliefs.

PURPOSE

Even though self-efficacy is a well-studied construct, there still remains a lack of studies from science content courses. Studies such as Luera and Otto (2005) did not collect data from preservice teachers when they were taking a science course but, rather, collected data from a science-methods course. Others such as Morrell and Carroll (2003) did collect data from preservice teachers taking a science course, but only collected data from traditional lecture courses and did not include data from inquiry-based, content courses specifically designed for preservice teachers.

Within the geosciences, a few studies have assessed student self-efficacy when they were taking content courses, but a gap remains in understanding the effect of such courses on self-efficacy. Slater et al. (2008) collected data from an inquiry-based science course for preservice teachers, but the course only covered concepts in astronomy, and changes in self-efficacy were not reported. Gosselin et al. (2010) described a professional-development workshop for in-service teachers that included seven preservice teachers. It is not known whether those seven students had taken any science-content courses before the study. Students, reported on in Baldwin (2014), completed a traditional introductory geology course with a laboratory section specifically designed for preservice elementary teachers. Lastly, Posnanski (2007) investigated the effect of a reformed Earth Science content course on student self-efficacy, but the small sample size ($n = 49$) resulted in insufficient power to confidently detect the observed effect sizes. In addition, it is not known whether those students had completed any other undergraduate science-content courses, and it appears that the course used as the treatment in that study included 3 h of lecture and 2 h of laboratory. This is in contrast to the course taught in the current study, which was taught entirely in a laboratory setting, with minimal lecture. Although studies like these suggest that Earth Science content courses using reform-minded pedagogies do improve teacher self-efficacy, the literature is lacking in data to verify that claim, especially when compared with the numerous studies that consistently demonstrate the relationship between science teaching-methods courses and gains in science-teaching self-efficacy. Furthermore, a gap exists concerning the effect that other undergraduate science content courses may have on student self-efficacy. Put another way, do students who have taken several undergraduate science content courses show the same gains in self-efficacy as students who have completed few or no science courses? The purpose of this study was to fill this gap by collecting self-efficacy data within an undergraduate course and comparing the results to the number of science courses that the students had previously competed.

RESEARCH METHODS

Study Context

Since the late 1980s, the University of Northern Iowa has offered science-content courses specifically designed for elementary education majors. These courses have evolved from activity-based to inquiry-based curricula that emphasize process skills and best teaching practices. These are not methods courses but do address best teaching practices by modeling hands-on, student-centered, science teaching from a constructivist perspective. We currently offer three such courses that cover content in life science (Inquiry into Life Science), physical science (Inquiry into Physical Science), and Earth/Space science (Inquiry into Earth and Space Science [IiESS]). All elementary and early childhood majors are required to complete this three-course sequence. Initially, IiESS was an optional course to obtain a teaching license, but the state of Iowa has mandated that, starting in 2015, all applicants for a K–8 teaching license must include science content in physical, life, and Earth/space science (Iowa Board of Educational Examiners, 2011). Consequently, an Earth Science course, such as IiESS, is now required for all students majoring in elementary education and early childhood education. Because of this policy change, the class demographics shifted in the fall of 2013 from predominately freshmen and sophomores to predominately juniors and seniors. To meet this increased demand for Earth Science courses, elementary education and early childhood education majors can also take one of the three lecture-based introductory Earth Science courses offered by the Earth Science Department; however, few students have taken those courses.

Each section of IiESS meets for 5 h/wk, and enrollment is capped at 28 students to ensure significant teacher–student interactions. Lectures are kept to a minimum, and students work with in collaborative groups to complete a series of conceptually based, guided-inquiry projects and activities. For example, in the unit on groundwater, the students develop a class-wide protocol for calculating the porosity of various Earth materials (sand, gravel, sand + gravel, and clay), make predictions for each material, and develop a protocol to determine the porosity of each material. When the sand and gravel yielded similar results, students developed an explanation for that discrepant event and discussed their explanations with the class as a whole. As a follow-up activity, the students were presented with a scenario of a contaminated monitoring well to reinforce the concept of rapid groundwater flow via poorly consolidated sediment. Through such activities, students personally interacted with a variety of Earth and Space concepts that, in many cases, involved analysis of real-world data. Such an approach was similar to that described by Posnanski (2007) but did not include an open-ended inquiry component, such as the one used by Avery et al. (2012). Even though IiESS is not a methods course, the lessons in the course modeled teaching practices that are consistent with recent reform initiatives, such as the NGSS (NGSS Lead States, 2013), and connections between this course and K–12 education were regularly made. Such an approach aligns with calls to support science-teacher self-efficacy by including constructivist pedagogies in content courses (Schoon and Boone, 1998; Bleicher and Lindgren, 2005; Liang and Richardson, 2009) and supports the goals taught in the science methods courses.

The subjects covered in this course were originally aligned with Iowa's core standards for K–5 elementary science classes (Iowa Department of Education, 2012) and included material from astronomy, geology, and meteorology. With Iowa's recent adoption of the NGSS, we are in the process of revising IiESS to align with those new standards. Students interested in teaching middle-school or upper-elementary science are encouraged to earn a basic science minor, which includes a second Earth and Space Science content course that expands upon the concepts taught in IiESS and introduces new concepts more typical of a middle-school or high-school course.

Research Questions

Given the predominance of studies on science-teaching self-efficacy that report data from science methods courses, and the paucity of data from science content courses, this study sought to answer the following research questions within the context of IiESS.

1. Does the science-teaching self-efficacy of preservice elementary teachers improve after taking an inquiry-based Earth Science content course?
2. Does science-teaching self-efficacy significantly vary according to the student's undergraduate year (freshman, sophomore, etc.) after taking an inquiry-based Earth Science content course?
3. Does the number of undergraduate science content courses affect the science-teaching self-efficacy of preservice elementary teachers after taking an inquiry-based Earth Science content course?

Data Instruments

For this study, I used the Science Teaching Efficacy Belief Instrument (STEBI-B) (Enochs and Riggs, 1990; Riggs and Enochs, 1990) to measure science teaching self-efficacy. This instrument contains 23 statements using a five-point Likert scale, which ranges from *strongly disagree* to *strongly agree*. Each statement asks participants to rate their confidence toward various activities associated with teaching science. The STEBI-B is written for preservice teachers, so all responses are written in future tense, and results are scored by assigning one point for a *strongly disagree* response to five points for *strongly agree* and then adding the scores together to obtain a total self-efficacy score that ranges from a low of 23 to a high of 115. Ten of the 23 questions are worded such that a person with a high level of self-efficacy would select *strongly disagree*, so answers to those questions must be reverse-coded before scoring.

The STEBI-B is divided into two subscales. (See Table I for sample questions from each subscale.) Thirteen questions assess Personal Science Teaching Efficacy (PSTE), which is the belief in one's ability to teach science. Possible scores on the PSTE scale range from a minimum of 13 to a maximum of 65. In most studies, pretest scores are in the mid- to upper-40s (Morrell and Carroll, 2003; Posnanski, 2007; Liang and Richardson, 2009), and posttest scores are in the low- to mid-50s (Bleicher, 2009; Bautista, 2011), although several studies have reported posttest scores in the high 40s (Watters and Ginns, 1995; Tosun, 2000b; Finson, 2001). The remaining 10 questions measure Science Teaching Outcome Expectancy (STOE), which is the understanding that good teaching practices directly influence student learning.

TABLE I: Sample questions from the STEBI-B arranged by subscale.

Personal Science-Teaching Efficacy
I know the steps necessary to teach science concepts effectively
I will typically be able to answer students' science questions
When a student has difficulty understanding a science concept, I will usually be at a loss as to how to help the student understand[1]
Science Teaching Outcomes Expectancy
The teacher is generally responsible for the achievement of students in science
When a student does better than usual in science, it is often because the teacher exerted a little extra effort
Increased effort in science teaching produces little change in students' science achievement

[1]Note: Questions in *italics* are examples of items that must be reverse coded before scoring.

Possible scores on the STOE subscale range from a minimum of 10 to a maximum of 50 points and published results typically range from the low to upper 30s (Watters and Ginns, 1995; Bleicher and Lindgren, 2005; Bautista, 2011) with a few studies reporting scores in the 20s (Cantrell et al., 2003; Mulholland et al., 2004) or 40s (Lumpe et al., 2000; Gosselin et al., 2010).

Studies using the STEBI-B consistently report significant gains in PSTE for students taking science teaching-methods courses (Huinker and Madison, 1997; Bleicher and Lindgren, 2005; Bautista, 2011) and inquiry-based content courses (Luera and Otto, 2005; Posnanski, 2007; Palmer et al., 2015) but not from lecture-based, science-content courses (Morrell and Carroll, 2003; Posnanski, 2007). Morrell and Carroll (2003) also noted that results from the STEBI-B might contain a ceiling effect because students with high PSTE pretest scores have less room for improvement than do students who score low on the pretest. In contrast, results from the STOE subscale are mixed. Several studies have found no significant improvement after completing a science content course (Utley et al., 2005; Bleicher, 2009; Liang and Richardson, 2009). Those studies that have reported significant gains in STOE have only come from science methods courses (Finson, 2001; Bautista, 2011). Luera and Otto (2005) have noted that STOE is resistant to change and suggest that factors beyond science content knowledge affect that construct.

Both subscales were demonstrated to be valid and reliable by Enochs and Riggs (1990) with Cronbach's-α of 0.90 (PSTE) and 0.76 (STOE). Subsequent studies have reported similar values (Bleicher and Lindgren, 2005; Bayraktar, 2011; Palmer et al., 2015). Bleicher (2004) reanalyzed the STEBI-B and found that rewording two questions improved the overall reliability of the STOE subscale. That revised instrument was used in this study. The data from this study were consistent with published data for PSTE (0.80) but was low for STOE (0.63).

Since its introduction, the STEBI-B has been widely used to measure preservice science teacher self-efficacy (Ginns et al., 1995; Cakiroglu et al., 2005; Schoon and Boone, 1998; Tosun, 2000a; Gencer and Cakiroglu, 2007; Hechter, 2011). It has also been successfully implemented in a variety of international settings, including Australia (Mulholland and Wallace, 1996, 2001; Mulholland et al., 2004), South Africa (Kiviet and Mji, 2003; Mji and Kiviet, 2003), Turkey (Tekkaya et al., 2004; Cakiroglu et al., 2005; Gencer and Cakiroglu, 2007), Denmark (Andersen et al., 2003), Jordan (Wesah, 2012), and Singapore (Ng, 2011). In all cases, the STEBI-B has proven to be a valid and reliable instrument across these varied populations and has yielded similar results; however, Cakiroglu et al. (2005) noted that cultural differences might explain observed differences in STEBI-B scores. These findings suggest that the STEBI-B is a robust instrument with a high degree of validity and reliability that can be applied in a variety of settings.

Data Collection and Analysis

Between fall 2011 and spring 2014, students enrolled in IiESS completed the STEBI-B during the first and final weeks of each semester. Besides data on self-efficacy, the students provided demographic data and informed consent. The results were tabulated into a spreadsheet and analyzed using the SPSS v.21 statistical software package (IBM Corporation, Armonk, NY).

Given that some studies have identified a positive relationship between the number of undergraduate science courses taken and science-teaching self-efficacy (Ramey-Gassert et al., 1996; Bleicher, 2004; Luera and Otto, 2005), it is possible that any observed gains in self-efficacy were due to differences in the number of undergraduate science courses completed by the students before taking IiESS. To address that issue, the students listed the number of undergraduate science courses they had completed before that semester, and those data were included in the overall analysis.

Furthermore, it could be argued that gains in self-efficacy were generated by experiences in teacher education courses, such as a science methods course and not from experiences in IiESS. Elementary education majors at the University of Northern Iowa are required to complete a four-level professional sequence. Each level requires in-class field experiences that begin with basic observations (Level I), creating and teaching a lesson (Level II), developing and implementing a short unit (Level III), and student teaching (Level IV). Levels I–III require more than 80 h of in-class teaching experience, and students must have completed all required courses before they can become student teachers, and the Level III experience includes a week-long immersion experience. Given that program, students enrolled in IiESS would not yet have begun their student-teaching experiences. Typically, students complete Level I during their freshman or sophomore year and Level III during their senior year, just one semester before student teaching. Level III courses include both a science methods course and a week-long immersive teaching experience, during which, the students are absent from any other university courses (such as IiESS). Direct data on which level each student had completed would have led to a more-robust study; however, it should be noted that few students (typically zero but no more than two) completed their Level III field experience each semester. Given the lack of students with a week-long absence, it is very likely that few students in this study had completed a methods course. Therefore, student undergraduate year (freshman, sophomore, etc.) was used as a proxy for progress through the teacher education program.

Data analysis consisted of within-subjects, repeated-measures, mixed analysis of variance (ANOVA) (Stevens, 2009), with the number of undergraduate science courses as the between-subject, independent variable and scores on the two STEBI-B subscales as the dependent variables. Within-subject designs combine the data collected multiple times from all participants rather than analyzing the differences between a pretest and posttest. Such an approach is preferable to a one-way ANOVA because participant self-efficacy is measured more than once, which reduces measurement error (Stevens, 2009). The repeated-measures design treats time (the pretest and posttest answers to the STEBI-B) and individual students as independent variables (Weinfurt, 2000), thus, providing a method for assessing growth in self-efficacy for the entire class rather than comparing self-efficacy scores between two or more groups. The results reported below consist of within-subjects analyses that assessed whether scores on the STEBI-B varied with student undergraduate year, and subsequent between-subjects analyses then assessed whether students from different undergraduate years yielded different scores on the STEBI-B. Additional post hoc analyses further assessed whether each measure varied between the pretest and posttest. With all of these analyses, the level of significance was set at $p = 0.05$.

To account for the Morrell and Carrol (2003) ceiling effect, individual student normalized-gain scores (Hake, 1998) were evaluated against both independent variables using a one-way ANOVA. Many authors calculate those normalized gain scores from course- or group-level data and use the results to describe the magnitude of change between student groups (Hake, 2002; Lunsford, 2004; Kortz et al., 2008). Others have conducted inferential statistics on normalized gains calculated for individual students (Cheng et al., 2004; Coletta and Phillips, 2005; Knight and Wood, 2005) Recently, Marx and Cummings (2007) adjusted Hake's formula to minimize negative outliers in the data by calculating negative gains differently than positive gains. This new metric (called normalized change) is a nonlinear function and, therefore, cannot be compared using inferential statistics. Given the need to compare normalized gains among groups of students, Hake's original formula was used.

The results from this study also contained sufficient power to detect medium to large effect sizes. Cohen (1988) defined the effect size for a t-test (called the d-value) as the mean difference divided by the standard deviation of the sample and provided estimates for values of d corresponding to small (0.20), medium (0.50), and large (0.80) effect sizes. Thus, d measures the number of standard deviations that separate the two means. As reported below, the 180 students included in this study provided sufficient data for analysis, and the tables provided by Cohen indicate that setting the power at 0.80, with a sample size of 180 students, allowed detection of effect sizes larger than $d = 0.29$.

RESULTS

Demographics

Between fall 2011 and spring 2014, 189 students enrolled in eight sections of Inquiry into Earth and Space Science

TABLE II: Number of participants in each university year.[1]

Semester	Section	n	University Year			
			Freshman	Sophomore	Junior	Senior
Fall 2011	1	26	15	5	2	4
	2	25	14	5	3	3
Spring 2012	1	22	3	15	4	0
Fall 2012	1	26	13	3	9	1
	2	20	8	2	7	3
Fall 2013	1	21	2	5	8	6
	2	21	3	3	9	6
Spring 2014	1	19	6	5	3	5
Totals		180	64	43	45	28

[1]Note: Data from Spring 2013 not available.

(IiESS), consented to participate in this study and provided all required data. Nine of those students were excluded from the study, even though they provided a complete suite of data. One student had already graduated with a bachelor's degree, and six students had completed more than two undergraduate science courses. Those seven students were demographic outliers and provided an insufficient sample size to create an additional category for any subsequent analysis. An inspection of the normalized gain scores revealed that two students had outlying gain scores of −275 and −300, each of which were more than 10 z-scores below the mean, whereas the data from the remaining students produced gains that were less than 4 z-scores from the mean. Very low normalized gain scores occur when a student scores very high on the pretest but much lower on the posttest. Including the two students with anomalous normalized gain scores changed the significance of all subsequent analyses, and they were, therefore, removed from the study. However, removing any additional students did not noticeably alter the final results, so only students with normalized gain scores less than −100 were removed. The resulting data pool consisted of 180 students from nine sections (Table II), who were predominately female (92%) and less than 24 y old (98%). Most (60%) were freshmen or sophomores; however, in fall 2013, the number of juniors and seniors were higher than their younger counterparts (Table II). This shift in demographics corresponded to an increased demand for the course generated by changing state licensure requirements (Cervato et al., 2013).

For the number of previous science courses, most freshmen and sophomores had completed few courses, whereas juniors and seniors had completed at least one course. Almost half of the students in the study (42%) had never taken an undergraduate science course, but 72% of those students were freshmen and 82% were freshmen or sophomores. In contrast, only 22% of the students had completed one science course (65% were freshmen or sophomores) and 36% had completed two courses. Of the 64 students who had competed two science courses, 18 were sophomores and one was a freshman. The exact nature of those other science courses is not known, but given that all elementary education majors were required to complete both Inquiry into Life Science and Inquiry into Physical Science, it is likely that these other courses were inquiry-based content courses designed for this student population rather than traditional lecture courses. However, it cannot be ruled out that some of the students in this study completed a lecture-based course from another institution. This conclusion is also consistent with anecdotal observations made by instructors from all three inquiry courses.

None of the 180 students in this study provided data on their progress through the teacher education program; however, the 61 students from the 2013–2014 academic year did provide such information. Those students were predominately juniors or seniors (61%) who had taken two science courses (53%). However, only one student had completed a science methods course before taking IiESS, and two students took their methods course concurrently with IiESS. This suggests that, unlike most studies that used the STEBI-B, the students in this study had not completed a science methods course. Even though these students represent only 34% of the sample population, this trend is consistent with observations from the previous sections, especially given that earlier sections had higher percentages of freshmen and sophomore students.

Other demographic variables, such as sex, parent education, or high school size, did not yield significant differences across any of the study's dependent variables, and no systematic differences were observed between those students included in this study and those that were excluded. In addition, preliminary repeated-measures, mixed ANOVAs found no significant difference between course section and PSTE [$F_{(7,172)} = 1.15$, $p = 0.332$] or STOE [$F_{(7,172)} = 0.88$, $p = 0.527$]. The lack of any significant differences between sections suggested that the data were uniform over the length of the study and could be aggregated for all subsequent analyses.

Overall Changes in Self-Efficacy

Before comparing the results from the STEBI-B to the two research questions, the results were analyzed for all 180 participants combined (Table III). Mean student PSTE scores were between 29 and 61 on the pretest and were 32–64 on the posttest. A PSTE score of 39 represented the midrange score on this scale and corresponded to answering all questions with the midpoint *not sure* value. Very few participants scored below that threshold (pretest = 29, posttest = 11), with five students scoring low on both occasions. The mean increase of 4.0 points was significant, with a moderate effect size [$t_{(1,\ 179)} = 10.96$, $p < 0.001$, $d =$

TABLE III: Mean pretest, posttest, and normalized gain values for PSTE and STOE compared with student undergraduate year.

	Pretest		Posttest		Paired *t*-Test[1]		Normalized Gain[2]	
Undergraduate Year	Mean	SD	Mean	SD	*p*	*d*	Mean	SD
Personal Science Teaching Efficacy (PSTE)								
Freshman	44.8	6.6	49.9	4.9	**<0.001**	0.88	20.8	25.0
Sophomore	45.4	5.9	49.2	5.1	**<0.001**	0.68	16.9	24.7
Junior	46.4	6.6	49.6	6.6	**<0.001**	0.49	16.1	27.0
Senior	46.0	7.0	49.3	7.7	**0.002**	0.46	17.2	29.3
All students	45.6	6.5	49.6	5.9	**<0.001**	0.65	18.1	26.0
Science Teaching Outcome Expectancy (STOE)								
Freshman	37.1	3.7	37.5	3.5	0.353	—	−1.1	26.3
Sophomore	35.0	3.4	35.1	4.4	0.837	—	0.5	19.6
Junior	34.8	3.6	35.9	3.9	**0.021**	0.30	6.3	20.4
Senior	34.9	5.1	37.1	4.9	**0.001**	0.45	13.0	21.6
All students	35.7	4.0	36.5	4.2	**0.001**	0.18	3.3	23.0

[1]Statistically significant p-values are in bold ($p < 0.05$).
[2]Normalized-gain values calculated using the formula from Hake (1998).

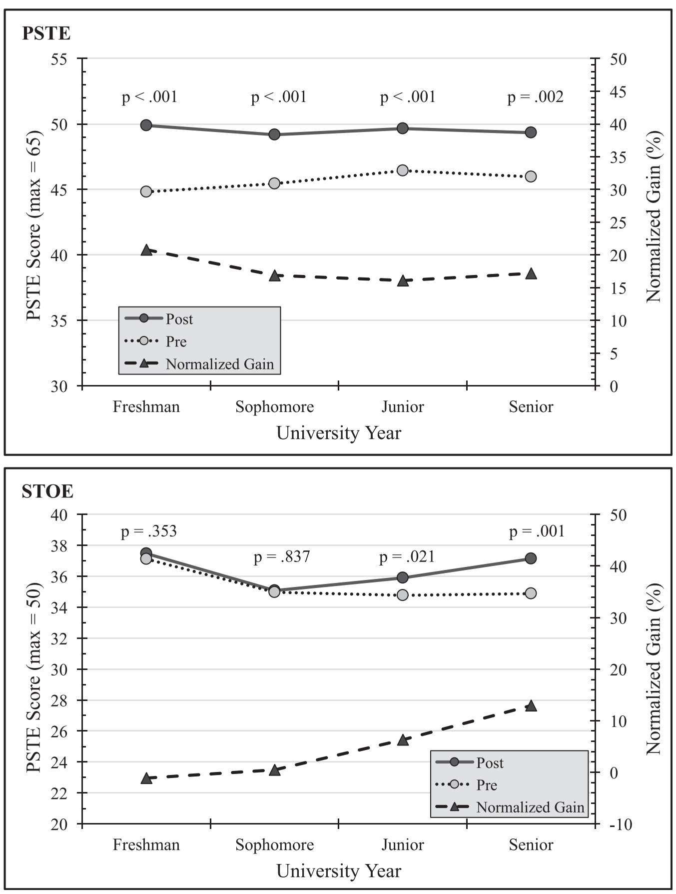

FIGURE 1: Mean pretest, posttest, and normalized-gain scores for the STEBI-B compared with student undergraduate year. Normalized-gain scores used the formula from Hake (1998). Significance of the paired *t*-tests for each undergraduate year is indicated by the *p*-values.

0.65]. The associated normalized-gain scores (Hake, 1998) ranged from 83.3 to 75.0, with 33 students earning a negative score, and an overall increase of 18.1%.

For STOE, the scores ranged from 25 to 48 on the pretest and 27 to 49 on the posttest. Eight students scored below the midrange STOE score of 30 on the pretest, but 13 scored below that value on the posttest. Four of the students with low pretest scores also scored low on the posttest, and two additional students scored 30. Across all students, the STOE score only improved by 0.8 points, although a two-tailed, paired *t*-test found that small increase significant but with a very small effect size, which was below the detection limits of this study [$t_{(1,179)} = 3.33$. $p = 0.001$, $d = 0.19$]. The normalized-gain scores ranged from 52.2% to −87.5%, with 65 students earning negative scores, but overall, the student scores improved by 3.3%.

The two independent variables in this study (*university year* and *number of prior science courses*) significantly correlated ($p < 0.001$, $r = 0.600$) with one another, which suggests that both variables may be measuring progress through the university and teacher education system. The results from the STEBI-B were compared against both variables, but the results for *university year* provided a clearer understanding of the observed variations than did the *number of prior science courses*. For that reason, the data and conclusions for *university year* are presented below. The STEBI-B data and results compared against the *number of science courses* are presented in the supplemental materials (available in the online journal at http://dx.doi.org/10.5408/14-022s1).

Self-Efficacy and Undergraduate Year

The results comparing PSTE to student undergraduate year (Fig. 1 and Table III) indicate that all four groups of students improved their scores, with the freshmen and sophomores showing slightly larger gains than the juniors and seniors had. The within-subjects, repeated-measures, mixed ANOVA tested whether students who had been in school longer experienced different gains in PSTE and produced no significant difference among students from different undergraduate years [$F_{(3,176)} = 1.60$, $p = 0.192$]. The between-subjects analysis compared PSTE values among the different years of schooling and also found no significant differences [$F_{(3,176)} = 0.17$, $p = 0.918$]. Like the results for the number of science courses, there was a significant difference between the pretest and posttest PSTE scores [$F_{(1,176)} = 101.96$, $p < 0.001$, $\eta^2_{partial}=0.367$], and paired t-tests indicated that students from all grade levels produced significant gains. The effect sizes for these gains decreased with time in school, with the freshmen posting the largest gain and effect size and the seniors showing the smallest gain and effect size. The normalized data yielded a similar trend. The freshmen had the largest raw and normalized gains, and a one-way ANOVA showed that the normalized gains were not significantly different among students from different undergraduate years [$F_{(3,176)} = 0.36$, $p = 0.784$].

The STOE data (Fig. 1 and Table III) revealed that the freshmen and sophomores had minimal increases in outcome expectancy with sophomores posting lower pretest and posttest scores than the freshmen. By contrast, the juniors and seniors gave similar pretest scores, but produced higher posttest scores. In addition, the pretest scores were highest for the freshmen and declined to a near constant level for the sophomores, juniors, and seniors. The posttest scores were also high for the freshmen and lower for the sophomores, but scores improved for the juniors and seniors. Both the within-subjects [$F_{(1,176)} = 3.440$, $p = 0.018$, $\eta^2_{partial}=0.055$] and between-subjects [$F_{(1,176)} = 4.152$, $p = 0.007$, $\eta^2_{partial}=0.066$] analyses found significant differences in STOE scores among the four grade levels. Similarly, the pretest and posttest values were significantly different [$F_{(1,176)} = 15.918$, $p < 0.001$, $\eta^2_{partial}=0.083$], but unlike PSTE, not all students showed significant gains. Paired t-tests demonstrated that outcome expectancy did not significantly improve for either the freshmen or sophomores but did significantly improve for the juniors and seniors, with the seniors posting a larger effect size than the juniors had (Table III). In addition, pairwise Bonferroni post hoc analyses for the combined pretest and posttest scores found that the freshmen scored significantly higher than the sophomores ($p = 0.012$) and juniors ($p = 0.040$) did. The normalized gains clearly show that trend, with the freshmen scoring the lowest (a negative score) and the seniors scoring the highest. A one-way ANOVA detected a significant difference [$F_{(3,176)} = 2.99$, $p = 0.033$] in those normalized-gain scores, but the Bonferroni post hoc analyses found that the scores among students from different years were not significantly different, except when comparing the freshmen to the seniors ($p = 0.041$, $d = 0.58$).

DISCUSSION

The results from this study indicate that taking an inquiry-based, Earth Science content course can improve both student PSTE and STOE; however, the magnitude of that improvement varies as students progressed through their undergraduate years. As noted above, the correlation between university year and number of science courses strongly suggests that both variables measured changes in the same construct. Therefore, the following discussion centers on the relationship between university year and both PSTE and STOE; a brief discussion on the number of science courses taken is available in the supplemental materials (available in the online journal at http://dx.doi.org/10.5408/14-022s1).

Overall Changes in Self-Efficacy

The significant increases indicated by the paired t-tests and between-subjects analyses indicate that students taking the IiESS course improved both their science-teaching self-efficacy (PSTE) and their outcome expectancy (STOE). The moderate effect size for PSTE and double-digit normalized gain suggests that the student self-efficacy improved while taking the course, but the small effect size and the normalized gain for outcome expectancy suggest that students experienced little or no gain in that construct. As discussed below, this lack of meaningful significance may be driven by the observed differences between freshmen and seniors.

Changes in PSTE

The significant increases in PSTE scores from pretest to posttest indicate that these students left the course with greater confidence about teaching science but the lack of a significant difference in the within-subjects analysis indicates that those differences were uniform regardless of one's year in school. The significant differences in the paired t-tests and moderate to large effect sizes suggest that the observed gains in self-efficacy did, in fact, occur and were not an artifact of the instrument. Those gains were similar in magnitude to published results from upper-level science-methods courses and indicate that the students taking the IiESS course made similar gains to upper classmen taking a science-methods course. In addition, the fact that these gains were similar across all four university years suggests that the observed increase came from student experiences in this course, rather than from their teacher-education courses (Huinker and Madison, 1997; Morrell and Carroll, 2003; Moseley and Utley, 2006). That was especially true for those freshmen who took IiESS during the fall of their freshmen year. However, the scores were lower in magnitude than many published results, which often report mean posttest scores in the low 50s (Liang and Richardson, 2009; Settlage et al., 2009; Bautista, 2011). The lower scores from this study most likely stem from the fact that many of the students were freshmen or sophomores and had just begun their teacher education program. The normalized gains further support this conclusion. The lack of any significant difference between different years indicates that the students from all levels experienced similar gains. This suggests that the improved PSTE values were a result of this course and not a result of experiences in methods or teaching field experiences. Even though the exact number of teaching experiences for each students is not known, it is highly unlikely that the freshmen had completed any experiences, and it is highly likely that the seniors had started the teacher-education courses. The nonsignificant results for both the raw data and the normalized gains further validate the conclusion that students of all ages improved their science self-efficacy.

As discussed below, freshmen had the greatest gains in PSTE, so it is possible that the observed differences in effect size were due to factors related to the freshman experience. More research is needed in this area to determine the contribution made by taking additional science courses to student PSTE.

Changes in STOE

Unlike the data for PSTE, not all students taking an inquiry-based content course improved their STOE scores. The minimal gains by freshmen and sophomores contrasted with the significant gains by the juniors and seniors. The gains by the upperclassmen contrasted with the results from other studies that found no gain in outcome expectancy after completing a science content course (Posnanski, 2007; Liang and Richardson, 2009; Baldwin, 2014) but aligned with studies that assessed juniors and seniors taking a science methods course (Finson, 2001; Bleicher and Lindgren, 2005; Bautista, 2011). The observed trend is best explained by the experiences those students brought to IiESS. Gains in outcome expectancy are generally attributed to direct teaching experiences within a school setting (Ginns et al., 1995; Cantrell et al., 2003; Bleicher, 2007). The minimal raw gains by the freshmen and sophomores may reflect their lack of experience teaching in an actual classroom, whereas the juniors and seniors had multiple opportunities in their Level I and Level II courses to experience the complex dynamics of teaching. This is consistent with the known demographics of the participants. Of the 61 students with known levels with the teacher education program, 42% of the freshmen and sophomores (10 of 24) had not yet taken any education courses, whereas only 5% (2 out of 37) of the juniors and seniors had not yet started that program. Similarly, only one freshman or sophomore had completed his or her Level II or III field experiences, but 22 of the juniors and seniors (59%) had done so. Even though these numbers represent only one-third of the total sample, they yielded the same results as the larger sample and support the claim that few freshmen and sophomores had completed more than their Level I observations.

The freshmen had larger pretest scores than the other groups had and larger posttest scores than all but the seniors. These findings are similar to those of Aydin and Boz (2010), who suggested that freshmen had not yet experienced the difficulties of teaching. It is likely that a similar dynamic occurred with the students taking IiESS. The freshmen (especially those in the fall semesters) had not yet taken any education courses, so their responses were more likely dictated by their youthful exuberance over starting a new chapter in their lives, rather than an accurate estimation of the teaching abilities or what it takes to succeed as a teacher.

Using Normalized Gain to Assess Self-Efficacy Beliefs

This study used the Hake (1998) normalized-change metric to account for possible ceiling effects (Morrell and Carroll, 2003), and the results accurately illustrated the changes in student self-efficacy and outcome expectancy. However, caution should be used when applying normalized change to student-attitude data, such as the STEBI. First, the formula for normalized change can generate negative outliers. If care is not taken, including these values will alter the conclusions from the study. In this investigation, a cutoff threshold of -100 was selected because that roughly corresponded to four standard deviations from the mean; however, data from other studies might find that that value too low. Using the Marx and Cummings (2007) normalized change metric would solve that problem, but the discontinuous nature of their formula limits its use. The second reason for caution relates to the original intent of the formula. Most studies use normalized change to evaluate how well their students learned key course concepts (e.g., Cheng et al., 2004; Kortz et al., 2008), and those studies often use a standardized instrument, such as the Force Concept Inventory (FCI) (Hestenes et al., 1992). In this situation, most student scores would be expected to improve; after all, that is the purpose and goal of teaching. Coletta and Phillips (2005) found that fewer than 2% of the 2,735 students they assessed produced negative normalized gain scores. By contrast, 33 students (18.3%) had negative normalized PSTE gains and 65 students (36.1%) had negative normalized STOE gains in this study. The high percentage of negative scores was likely due to the nature of the STEBI. Studies on the use of personal response systems in science lecture courses have consistently found that 5%–6% of students give incorrect responses after initially providing correct answers (Crouch et al., 2001; Steer et al., 2009). They attributed those decreases in learning to guessing or random data entry. The few negative scores on such instruments as the FCI probably had a similar source; however, unlike the FCI, the STEBI measures student beliefs about their own teaching ability, so negative values cannot be explained by random error. After all, there are no incorrect responses on a belief instrument, so the negative gain scores indicate a decrease in a student's belief in his or her ability to teach. In this case, it may be likely, although uncommon, that students experienced a large decline in self-efficacy over the course of a semester (especially for freshmen or sophomores). Therefore, using normalized gain to evaluate data from instruments like the STEBI will likely result in a large range of scores, including negative scores, which can unduly influence any inferential statistics. The lesson learned is that normalized-gain scores can be used in this context but with caution and attention to the magnitude of negative values.

Sources of Observed Changes in Self-Efficacy

Prior studies from methods courses have credited positive teaching experiences (mastery experiences) for observed increase in self-efficacy (Posnanski, 2007; Bautista, 2011). It is unlikely that is the reason for the increases in PSTE observed in this study. Many of these students were freshmen and sophomores, who were just beginning their teacher education courses and had very few in-class teaching experiences. In fact, for the fall semesters, the freshmen were just a few months removed from high school, so it seems unlikely that those students would have engaged in significant classroom teaching over the course of their first semester on campus. It is far more likely that the students improved their science-teaching self-efficacy because they better understood concepts they had not learned in high school. This conclusion agrees with the Stone (2016) qualitative study that found that students taking the IiESS course attributed their improved self-efficacy to gains in their Earth Science content knowledge. It is likely that those students entered the course knowing very little about Earth Science topics and left the course with a better understanding of geology, meteorology, and astronomy.

The Palmer (2006b) cognitive content mastery theory is a more likely source for the observed increases in PSTE. From that perspective, the students improved their self-efficacy through positive learning experiences of new content material rather than from having positive experiences mastering the art of teaching, which may be especially true for an Earth Science course such as IiESS because students often enter the course with little or no formal training in Earth Science (Tosun, 2000b). Stone (2016) found that those students often cited the newfound knowledge as a major reason for improving their self-efficacy. It is possible that the Earth and Space subject matter covered in this course was new to the students, so they may have derived greater perceived satisfaction from the course, which led to significant increases in self-efficacy.

IMPLICATIONS FOR TEACHING

The results from this study have several implications related directly to preparing preservice elementary teachers. First, the findings from this study align with comments by Wenner (2001) and Bender (2007) that preservice elementary teachers greatly benefit from science content courses that use a constructivist/inquiry pedagogy rather than traditional lectures. This case is further strengthened by studies from methods courses that include a significant content component (Palmer, 2006b; Liang and Richardson, 2009), and others who have called for content courses to explicitly address issues related to self-efficacy (Schoon and Boone, 1998) and pedagogy (Hechter, 2011). Geoscience departments should consider offering inquiry-based content courses designed explicitly for preservice elementary teachers. Only in such courses can instructors weave issues such as pedagogy and self-efficacy with content material. Doing so not only improves the content knowledge of these future teachers but also makes them better prepared to teach science in the classroom.

A second implication from this study is the link between science content courses for preservice elementary teachers and adoption of reform-minded pedagogies. Numerous studies have documented that people teach the way they were taught, which is often via lecture, confirmatory laboratory courses, and "busy work" (Stuart and Thurlow, 2000; Phelps and Lee, 2003; Choi and Ramsey, 2009). These attitudes are carried into teacher-education programs and beyond. Breaking that cycle requires multiple experiences with inquiry teaching beyond a well-planned methods course. Content courses such as IiESS can support methods courses by providing an additional layer of experiences in which the students learn firsthand the utility of this pedagogy (Weld and Funk, 2005). By embedding sound teaching practices within a science-content course, preservice teachers are shown both the effectiveness of inquiry teaching as well as the methods needed to teach in that manner. Subsequent methods courses can then build on the context provided by these earlier experiences.

This study has demonstrated that significant growth in student self-efficacy toward teaching science can be fostered in content courses as well as in methods courses. In addition, the results from this study suggest that taking multiple science courses does not significantly improve self-efficacy. Yet this study took place within the context of an Earth Science course, so further research is needed to clarify whether similar gains in self-efficacy would occur in a physical or life science course or whether students make incremental gains after taking more than one course.

Acknowledgments

I would like to thank the Northern Iowa Department of Earth Science and Provost's Office for partial support of this project. This article greatly benefitted from comments by three anonymous reviewers whose suggestions greatly improved the final manuscript. I would also like to thank Dawn Del Carlo, Jill Maroo, Jeff Morgan, and Sarah Boesdorfer, for helpful suggestions on an earlier draft, and Jacqueline Treanor for help with data input and analysis.

REFERENCES

Andersen, A.M., Dragsted, S., Evans, R.H., and Sorensen, H. 2003. Transforming the standard instrument for assessing science teacher's self-efficacy Beliefs (STEBI) for use in Denmark. *In* Psillos, D., Kariotoglou, P., Tselfes, V., Hatzikrantiotis, E., Fassoulopoulos, G., and Kallery, M., eds., Science education research in the knowledge-based society. Dordrecht, The Netherlands, Kluwer Academic Publishers, p. 395–403.

Avery, L.M., and Meyer, D.Z. 2012. Teaching science as science is practiced: Opportunities and limits for enhancing preservice elementary teachers' self-efficacy for science and science teaching. *School Science and Mathematics,* 112(7):395–409.

Aydin, S., and Boz, Y. 2010. Pre-service elementary science teachers' science teaching efficacy beliefs and their sources. *Elementary Education Online,* 9(2):694–704.

Baldwin, K.A. 2014. The science teaching self-efficacy of prospective elementary education majors enrolled in introductory geology lab sections. *School Science and Mathematics,* 114(5):206–213.

Bandura, A. 1977. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review,* 84(2):191–215.

Bandura, A. 1984. Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research,* 8(3):231–255.

Bandura, A. 1997. Self-efficacy—The exercise of control. New York: W.H. Freeman & Company, p 604.

Bautista, N. 2011. Investigating the use of vicarious and mastery experiences in influencing early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education,* 22(4):333–349.

Bayraktar, S. 2011. Turkish preservice primary school teachers' science teaching efficacy beliefs and attitudes toward science: The effect of a primary teacher education program. *School Science and Mathematics,* 111(3):83–92.

Bender, K. M. 2007. Everyday physics: Listening to pre-service teachers reflect on learning and teaching science through inquiry [M.S. thesis]. Akron, OH: University of Akron. p. 67.

Bleicher, R., 2004. Revisiting the STEBI-B: Measuring self-efficacy in preservice elementary teachers. *School Science and Mathematics,* 104(8):383–390.

Bleicher, R. 2007. Nurturing confidence in preservice elementary science teachers. *Journal of Science Teacher Education,* 18(6):841–860.

Bleicher, R. 2009. Variable relationships among different science learners in elementary science-methods courses. *International Journal of Science and Mathematics Education,* 7(2):293–313.

Bleicher, R., and Lindgren, J. 2005. Success in science learning and preservice science teaching self-efficacy. *Journal of Science Teacher Education,* 16(3):205–225.

Britner, S.L., and Pajares, F. 2006. Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching,* 43(5):485–499.

Brouwers, A., and Tomic, W. 2000. A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education,* 16(2):239–253.

Cakiroglu, J., Cakiroglu, E., and Boone, W.J. 2005. Pre-service teacher self-efficacy beliefs regarding science teaching: A comparison of pre-service teachers in Turkey and the USA. *Science Educator,* 14(1):31–40.

Cantrell, P., Young, S., and Moore, A. 2003. Factors affecting science teaching efficacy of preservice elementary teachers. *Journal of Science Teacher Education,* 14(3):177–192.

Caprara, G.V., Barbaranelli, C., Steca, P., and Malone, P.S. 2006. Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology,* 44(6):473–490.

Cervato, C., Kerton, C., Peer, A., Hassall, L., and Schmidt, A. 2013. The big crunch: A hybrid solution to Earth and space science instruction for elementary education majors. *Journal of Geoscience Education,* 61(2):173–186.

Cheng, K.K., Thacker, B.A., Cardenas, R.L., and Crouch, C. 2004. Using an online homework system enhances students' learning of physics concepts in an introductory physics course. *American Journal of Physics,* 72(11):1447–1453.

Choi, S., and Ramsey, J. 2009. Constructing elementary teachers' beliefs, attitudes, and practical knowledge through an inquiry-based elementary science course. *School Science and Mathematics,* 109(6):313–324.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. New York: Psychology Press. p. 567.

Coletta, V.P., and Phillips, J.A. 2005. Interpreting FCI scores: Normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics,* 73(12):1172–1182.

Crouch, C.H., Mazur, E., Beatty, I.D., Gerace, W.J., Leonard, W.J., and Dufresne, R.J. 2001. Peer instruction: Ten years of experience and results: *American Journal of Physics,* 69(9,):970–977.

Enochs, L.G., and Riggs, I.M. 1990. Further development of an elementary science teaching efficacy belief instrument: A preservice elementary scale. *School Science and Mathematics,* 90(8):694–706.

Finson, K. 2001. Investigating preservice elementary teachers' self-efficacy relative to self-image as a science teacher. *Journal of Elementary Science Education,* 13(1):31–41.

Gencer, A.S., and Cakiroglu, J. 2007. Turkish preservice science teachers' efficacy beliefs regarding science teaching and their beliefs about classroom management. *Teaching and Teacher Education,* 23(5):664–675.

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

Gibson, S., and Dembo, M.H. 1984. Teacher efficacy: A construct validation. *Journal of Educational Psychology,* 76(4):569–582.

Ginns, I.S., Tulip, D.F., Watters, J.J., and Lucas, K.B. 1995. Changes in preservice elementary teachers' sense of efficacy in teaching science. *School Science and Mathematics,* 95(8):394–400.

Gosselin, D.C., Thomas, J., Redmond, A., Larson-Miller, C., Yendra, S., Bonnstetter, R.J., and Slater, T.F. 2010. Laboratory earth: A model of online K–12 teacher coursework: *Journal of Geoscience Education,* 58(4):203–213.

Hake, R. 2002. Lessons from the physics education reform effort. *Ecology and Society,* 5(2):1–28.

Hake, R.R. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics,* 66:64–74.

Hechter, R. 2011. Changes in preservice elementary teachers' personal science teaching efficacy and science teaching outcome expectancies: The influence of context. *Journal of Science Teacher Education,* 22(2):187–202.

Hestenes, D., Wells, M., and Swackhammer, G. 1992. Force concept inventory. *Physics Teacher,* 30(3):141–151.

Huinker, D., and Madison, S. 1997. Preparing efficacious elementary teachers in science and mathematics: The influence of methods courses. *Journal of Science Teacher Education,* 8(2):107–126.

Iowa Board of Educational Examiners. 2011. IBEE Web site. Available at http://www.boee.iowa.gov/ (accessed 21 June 2013).

Iowa Department of Education. 2012. Iowa Core Science. Available at http://educateiowa.gov/index.php?option=com_content&view=article&id=2330&Itemid=4342 (accessed 21, June 2013).

Kiviet, A.M., and Mji, A. 2003. Sex differences in self-efficacy beliefs of elementary science teachers. *Psychological Reports,* 92(1):333–338.

Klassen, R., Tze, V., Betts, S., and Gordon, K. 2011. Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review,* 23(1):21–43.

Knight, J.K., and Wood, W.B. 2005. Teaching more by lecturing less. *Cell Biology Education,* 4(4):298–310.

Kortz, K., Smay, J., and Murray, D. 2008. Increasing learning in introductory geoscience courses using lecture tutorials. *Journal of Geoscience Education,* 56(3):280–290.

Liang, L., and Richardson, G. 2009. Enhancing prospective teachers' science teaching efficacy beliefs through scaffolded, student-directed inquiry. *Journal of Elementary Science Education,* 21(1):51–66.

Luera, G.R., and Otto, C.A. 2005. Development and evaluation of an inquiry-based elementary science teacher education program reflecting current reform movements. *Journal of Science Teacher Education,* 16(3):241–258.

Lumpe, A.T., Haney, J.J., and Czerniak, C.M. 2000. Assessing teachers' beliefs about their science teaching context. *Journal of Research in Science Teaching,* 37(3):275–292.

Lunsford, S.K. 2004. Integrated lecture and laboratory chemistry components of science education program for early and middle childhood education majors. *Journal of Chemical Education,* 81(5):685.

Marx, J.D., and Cummings, K. 2007. Normalized change. *American Journal of Physics,* 75(1):87–91.

Mji, A., and Kiviet, A.M. 2003. Psychometric characteristics of the science teaching efficacy belief inventory in South Africa. *Psychological Reports,* 92(1):325–332.

Morrell, P.D., and Carroll, J.B. 2003. An extended examination of preservice elementary teachers' science teaching self-efficacy. *School Science and Mathematics,* 103(5):246–251.

Moseley, C., and Utley, J. 2006. The effect of an integrated science and mathematics content-based course on science and mathematics teaching efficacy of preservice elementary teachers. *Journal of Elementary Science Education,* 18(2):1–12.

Muijs, D., and Reynolds, D. 2002. Teachers' beliefs and behaviors: What really matters? *Journal of Classroom Interaction,* 37(2):3–15.

Mulholland, J., Dorman, J.P., and Odgers, B.M. 2004. Assessment of science teaching efficacy of preservice teachers in an Australian university. *Journal of Science Teacher Education,* 15(4):313–331.

Mulholland, J., and Wallace, J. 1996. Breaking the cycle: Preparing elementary teachers to teach science. *Journal of Elementary Science Education,* 8(1):17–38.

Mulholland, J., and Wallace, J. 2001. Teacher induction and elementary science teaching: Enhancing self-efficacy. *Teaching and Teacher Education,* 17(2):243–261.

Ng, W.-L. 2011. A study of Singapore female primary teachers self-efficacy for teaching science [Ph.D. dissertation]. Durham, England: Durham University. p. 205.

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

Pajares, F. 1996. Self-efficacy beliefs in academic settings. *Review of Educational Research,* 66(4):543–578.

Palmer, D., 2006a. Durability of changes in self-efficacy of preservice primary teachers. *International Journal of Science Education,* 28(6):655–671.

Palmer, D. 2006b. Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education,* 36:337–353.

Palmer, D., Dixon, J., and Archer, J. 2015. Changes in science teaching self-efficacy among primary teacher education students. *Australian Journal of Teacher Education,* 40(12):27–40.

Phelps, A.J., and Lee, C. 2003. The power of practice: What students learn from how we teach. *Journal of Chemical Education,* 80(7):829.

Plourde, L.A. 2002. The influence of student teaching on preservice elementary teachers' science self-efficacy and outcome expectancy beliefs. *Journal of Instructional Psychology,* 29(4):245–253.

Posnanski, T.J. 2007. A redesigned geoscience content course's impact on science teaching self-efficacy beliefs. *Journal of Geoscience Education,* 55(2):152–157.

Ramey-Gassert, L., Shroyer, M.G., and Staver, J.R. 1996. A qualitative study of factors influencing science teaching self-efficacy of elementary level teachers. *Science Education,* 80(3):283–315.

Rice, D., and Roychoudhury, A. 2003. Preparing more confident preservice elementary science teachers: One elementary science methods teacher's self-study. J*ournal of Science Teacher Education,* 14(2):97–126.

Riggs, I.M., and Enochs, L.G. 1990. Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education,* 74(6):625–637.

Ross, J. A. 1992. Teacher efficacy and the effects of coaching on student achievement. *Canadian Journal of Education / Revue canadienne de l'éducation,* 17(1):51–65.

Schoon, K.J., and Boone, W.J. 1998. Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education,* 82(5):553–568.

Settlage, J., Southerland, S.A., Smith, L.K., and 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.

Slater, S.J., Slater, T.F., and Shaner, A. 2008. Impact of backwards faded scaffolding in an astronomy course for pre-service elementary teachers based on inquiry. *Journal of Geoscience Education,* 56(5):408–416.

Steer, D.N., McConnell, D.A., Gray, K., Kortz, K., and Liang, X. 2009. Analysis of student responses to peer-instruction conceptual questions answered using an electronic response system: Trends by gender and ethnicity. *Science Educator,* 18(2):30–38.

Stevens, J. P. 2009. Applied multivariate statistics for the social sciences. New York, Routledge. p. 651.

Stone, M. A. 2016. Analysis of the STEBI-B instrument as it is used to predict the science teaching self-efficacy beliefs of preservice elementary teachers [M.A. thesis]. Cedar Falls, IA: University of Northern Iowa. p. 90.

Stuart, C., and Thurlow, D. 2000. Making it their own: Preservice teachers' experiences, beliefs, and classroom practices. *Journal of Teacher Education,* 51(2):113.

Tekkaya, C., Cakiroglu, J., and Ozkan, O. 2004. Turkish pre-service science teachers' understanding of science and their confidence in teaching it. *Journal of Education for Teaching,* 30(1):57–68.

Tosun, T. 2000a. The beliefs of preservice elementary teachers toward science and science teaching. *School Science and Mathematics,* 100(7):374–379.

Tosun, T. 2000b. The impact of prior science course experience and achievement on the science teaching self-efficacy of preservice elementary teachers. *Journal of Elementary Science Education,* 12(2):21–31.

Tschannen-Moran, M., Hoy, A.W., and Hoy, W.K. 1998. Teacher efficacy: Its meaning and measure. *Review of Educational Research,* 68(2):202–248.

Usher, E.L., and Pajares, F. 2006. Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology,* 31(2):125–141.

Usher, E.L., and Pajares, F. 2008. Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research,* 78(4):751–796.

Utley, J., Moseley, C., and Bryant, R. 2005. Relationship between science and mathematics teaching efficacy of preservice elementary teachers. *School Science and Mathematics,* 105(2):82–87.

van Dinther, M., Dochy, F., and Segers, M. 2011. Factors affecting students' self-efficacy in higher education. *Educational Research Review,* 6(2):95–108.

Watters, J.J., and Ginns, I.S. 1995. Origins of, and changes in preservice teachers' science teaching self efficacy. Available at http://eprints.qut.edu.au/1761/1/1761.pdf (accessed 1 November 2016).

Weinfurt, K.P. 2000. Repeated measures analysis: ANOVA, MANOVA, and HLM. *In* Grimm, L.G., and Yarnold, P.R., ed. Reading and understanding more multivariate statistics. Washington, DC: American Psychological Association, p. 317–361.

Weld, J., and Funk, L. 2005. "I'm not the science type": Effect of an inquiry biology content course on preservice elementary teachers' intentions about teaching science. *Journal of Science Teacher Education,* 16(3):189–204.

Wenner, G. 2001. Science and mathematics efficacy beliefs held by practicing and prospective teachers: A 5-year perspective. *Journal of Science Education and Technology,* 10(2):181–187.

Wesah, H.A. 2012. Teaching efficacy and teaching performance among student teachers in a Jordanian childhood education program. *Journal of Early Childhood Teacher Education,* 33(2):163–177.

Wysession, M.E. 2012. Implications for earth and space in new K–12 science standards. *EOS,* 93(46):465–472.

Zundans-Fraser, L., and Lancaster, J. 2012. Enhancing the inclusive self-efficacy of preservice teachers through embedded course design. *Education Research International,* 2012:1–8.