International Journal of Science Education

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

A longitudinal investigation of the science teaching efficacy beliefs and science experiences of a cohort of preservice elementary teachers

James Deehan, Lena Danaia & David H. McKinnon

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
https://doi.org/10.1080/09500693.2017.1393706

A longitudinal investigation of the science teaching efficacy beliefs and science experiences of a cohort of preservice elementary teachers

James Deehan [a], Lena Danaia [a] and David H. McKinnon [b]

[a]Faculty of Arts and Education, Charles Sturt University, Bathurst, Australia; [b]School of Education, Edith Cowan University, Perth, Australia

ABSTRACT

This paper assesses the relationship between participation in two tertiary science courses and the science teaching efficacy beliefs (STEBs) of one cohort of preservice elementary teachers over a four-year period. Two Type II case studies were conducted within the courses. Data were collected through 26 administrations of the Science Teaching Efficacy Belief Instrument-B and semi structured interviews. Results showed that participation in the subjects covaried with increases in the participants' STEBs. These increases in STEBs remained durable for two years. Implications for these findings are discussed within the paper.

ARTICLE HISTORY
Received 20 October 2016
Accepted 14 October 2017

KEYWORDS
Science teaching efficacy; preservice elementary teachers; elementary science; preservice teacher education; longitudinal

Introduction

The negative science attitudes of many elementary teachers have been illustrated repeatedly in research literature for decades (de Laat & Watters, 1995; Mulholland, Dorman, & Odgers, 2004; Schibeci, 1984). Many teachers do not feel confident in their ability to teach science because of their inadequate science content knowledge (Palmer, 2011; Tytler, Smith, Grover, & Brown, 1999). Other factors that contribute to poor science attitudes include; low Pedagogical Content Knowledge (PCK), negative personal educational experiences and the perceived cognitive difficulty of science content matter (Jarrett, 1999; Juriševič, Glažar, Pučko, & Devetak, 2008; Nillson & Loughran, 2011; Skamp & Mueller, 2001).

Given the largely negative views of science held by elementary teachers, it is unsurprising that science curricula are often marginalised and distorted (e.g. Appleton, 2003; Appleton & Kindt, 2002). During the 1990s, it was often reported that science was taught for an hour a week in Australian elementary classrooms (Gough et al., 1998). Later reports indicated that the average amount of time spent in elementary science had decreased to 40 minutes per week (Goodrum & Rennie, 2007). Analysis of the 2015 Trends in International Mathematics and Science Study (TIMSS) data indicates that Australian elementary schools may be allocating just 5.6% of available curriculum time to Science education;

CONTACT James Deehan jdeehan@csu.edu.au Faculty of Arts and Education, Charles Sturt University, Bathurst, NSW, Australia

a finding which would rank a paltry 39th amongst the participating OECD countries (Martin, Mullis, Foy, & Hooper, 2016).

The limited science that is taught in elementary schools is often teacher centric and transmissive (Goodrum & Rennie, 2007; Jarrett, 1999; Smith, 2014; Weiss, 1994). According to Kelly (2000, p. 756),

> what typically transpires in the science classroom is not the hands on, minds on paradigm that demonstrates a fusion of pedagogical strategy and content knowledge. Rather, science teaching is often reduced to a collection of facts, discussions about assigned readings, and an occasional activity.

Indeed note taking, completing worksheets and teacher demonstrations are common place in elementary schools (Goodrum & Rennie, 2007; Goodrum, Hackling, & Rennie, 2001). In fact, many elementary teachers are unaware of what constitutes 'good' science teaching (Appleton, 2003). This may stem from conceptually poor understanding of science PCK that leads to inappropriate pedagogical selection. Teachers often incorrectly believe that physical manipulation alone gives students control over their learning.

The achievement level of Australian elementary students in science has declined as other nations have advanced (Gonski, 2011). The TIMSS assesses and compares the scientific content knowledge and scientific literacy of Year 4 students from as many as 52 nations (Thomson et al., 2012). Table 1 summarises Australia's elementary science performance in TIMSS from 1995 through 2015 (Gonski, 2011; Martin et al., 1997; Martin et al., 2016; Mullis, Martin, Gonzalez, & Chrostowski, 2004; Thomson et al., 2012; Thomson, Wernet, Underwood, & Nicholas, 2008). Despite consistently scoring above the OECD average, the performance of Australian Grade 4 students has declined since 1995 and more recent testing has shown stagnation. Australia's sharpest decline has occurred in the international rankings. Nations such as Italy, Slovakia, Hong Kong and Hungary now score more highly than Australia, with nations such as Japan, Finland and Russia producing more scientifically literate elementary school students. There is a large gap between high performing and low performing students in Australia. The above average students represent 27% of the total population, and they are reaching similar levels to Singaporean and Finnish students. The tail group is larger (36%) and these students equivalent to those from lower performing nations such as Armenia, Qatar and Oman. Such disparity is likely to be another inhibiting factor in the teaching of elementary science in Australia. Perhaps the most concerning finding is Australia's fall below the 'High threshold'. Per Thomson et al. (2012), a mean score of 550 or above would suggest that students are able to apply science knowledge and skills to novel situations beyond the classroom context. Australia has failed to reach this level for nearly two decades. This is an indicator that elementary science educators in Australia are broadly failing to develop the scientific literacy of their students.

Table 1. Australian year 4 students' science achievement in the TIMSS.

Year	Australia's Score	Mean Score	Ranking
1995	562	524	5th
2003	521	489	11th
2007	527	500	13th
2011	516	486	19th
2015	524	500	25th

Longitudinal professional development programmes, with ongoing support, can improve the science teaching attitudes and practices of inservice elementary teachers (Duran, Ballone-Duran, Haney, & Beltyukova, 2009; Lumpe, Czerniak, Haney, & Beltyukova, 2012). Tytler et al. (1999) suggested that a longitudinal approach to professional development covaries with improved science confidence and knowledge. Palmer (2011) implemented another longitudinal professional development programme, the results of which indicated that cognitive mastery and in situ feedback improved the science teaching efficacy of the teachers. While professional development opportunities are undoubtedly beneficial, there are considerable financial and time requirements to the provision of the embedded, longitudinal support needed to improve the science teaching efficacy and practices of inservice elementary teachers. The increasing mean age of inservice teachers may render interventions within preservice teacher education programmes more efficient (Harris & Farrell, 2007; NSW DE, 2015; NSW DEC, 2011).

A growing body of literature has shown that science interventions targeting preservice elementary teachers can successfully alter negative perceptions and address knowledge deficits (Cooper, Kenny, & Fraser, 2012; Watters & Ginns, 2000). Explicit instruction on the nature of science through the use of scaffolded argument has been shown to both improve preservice teachers' understanding of science and alleviate some of their existing negative views towards science (McDonald, 2010). Cooperative learning, in an authentic setting, can improve both science attitudes and science teaching efficacy (Watters & Ginns, 2000). Bleicher and Lindgren (2005) found that the use of constructivist approaches to address students' alternate scientific conceptions produced large effect size increases (Cohen's $d = 1.2$) in preservice teachers' personal science teaching efficacy beliefs (STEBs). Cooper et al. (2012) found that mentoring programmes between preservice and inservice teachers were beneficial for both parties. It should be noted that the negative trends within elementary science classrooms do not reflect the positive reports from the tertiary level. Amidst the promising results of tertiary science education programmes, there is still a lack of consistency in how science is taught to prospective elementary teachers (Palmer, 2007).

D. Palmer (2006) found that the science teaching efficacy gains made by a group of preservice elementary teachers as they participated in a cooperative, inquiry-based science subject remained durable for nine months after the course had been completed. Richardson and Liang (2008) considered the durability of science teaching efficacy by utilising the delay period between science course offerings. They found that the participants' STEBs increased in the absence of treatment. Ginns, Tulip, Watters, and Lucas (1995) used a longitudinal design to evaluate a four-year preservice elementary teacher education programme. They found that participants' STEBs did not improve during their undergraduate studies.

There is a clear disconnection between the reported problems within elementary science education, such as poor science teaching efficacy (Palmer, 2011), student disengagement (DeWitt, Archer, & Osborne, 2014), inadequate curriculum time (Tytler, Osborne, Williams, Tytler, & Clark, 2008) and diminishing achievement levels (Martin et al., 2016); and the positive science outcomes, such as improved efficacy (Sang et al., 2012), conceptual understandings (Menon & Sadler, 2016), and pedagogical content knowledge (Akerson, Pongsanon, Rogers, Carter, & Galindo, 2017), reported within preservice teacher education programmes. The research presented in this paper begins to explore this apparent disconnection via a four year, longitudinal design; which is a clear contribution to a research domain dominated by single course, pre-post-test designs (Deehan, 2016). The four-year data collection period allows for durability of the outcomes from a two-year science programme (SC108 and SC308) to be assessed for two years in the absence of formal science intervention as the participants proceed to graduation. Such depth is almost unprecedented within the existing literature.

The aims of this research are threefold. The first aim is to assess the relationship between preservice elementary teachers' participation in a complex, innovative science programme with two complementary courses (SC108 and SC308) and their STEBs. The second aim is to determine if any changes to participants' STEBs that occur within the science programme remain durable for up to two years in the absence of treatment. The third aim is to inductively explore the participants' attitudes towards and perceptions of science teaching as they progress through the science courses and beyond. The questions are as follows:

1. What is the relationship between preservice elementary teachers' participation in a complex, innovative science programme with two complementary courses (SC108 and SC308) and their STEBs?
2. If participants' STEBs change during the science programme (SC108 and SC308), do these changes remain durable for up to two years in the absence of treatment?
3. What were the participants' attitudes towards and perceptions of science teaching as they progressed through the science courses and beyond?

The significant contributions of the research presented in this paper can be summarised as follows: targeting a stakeholder group with the most potential for long-term, direct impact on the provision of high quality science education, providing a replicable model for a preservice elementary science education domain marred by inconsistency, and finally, assessing the durability of STEBs for a two-year period after the completion of a science programme. More generally, it could be argued that the need for research in this field has been heightened as forces such as globalisation and rapid technological advancement have increased the necessity for scientific skills and knowledge within the modern workforce (Levy & Murnane, 2006).

Theoretical framework

Reports of poor science content knowledge (Appleton, 1992, 2002, 2003; Howitt, 2007) and low PCK (Hechter, 2010) by elementary teachers may be evidence of low confidence and self-efficacy. Teachers cannot consider how their science teaching influences student outcomes until they themselves believe that their teaching practice will have the intended effects on student learning. At a basic level, self-efficacy can be defined as an individual's judgement of his or her competence to execute a task (Bandura, 1977, 1986). Self-efficacy is one of the strongest predictors of human motivation and behaviour (Bandura, 1986).

Teacher Efficacy can be defined as a 'teacher's belief or conviction that they can influence how well students learn, even those who may be difficult or unmotivated' (Guskey & Passaro, 1994, p. 4). It has also been acknowledged that teachers' self-efficacy beliefs interact with external factors beyond their immediate control; thus many contemporary definitions of teacher efficacy have been expanded to include both internal and external loci of control Teacher Efficacy has been found to correlate positively with desirable outcomes in both teachers and students (Goddard, Hoy, & Hoy, 2000). Modelled on Bandura's (1977, 1986) and Rotters' (1966) earlier works, the construct is comprised of Personal Teaching Efficacy (PTE) and General Teaching Efficacy (GTE). PTE describes an individual's belief in their own ability to overcome contextually specific factors to promote student learning (Coladarci, 1992; Gordon & Debus, 2002). GTE is the belief that student learning can be influenced by effective teaching (Enochs & Riggs, 1990; Gibson & Dembo, 1984). A teacher with high GTE beliefs is more likely to believe that effective teaching can override potentially detrimental social, economic and cultural factors to influence student learning positively (Tschannen-Moran & Hoy, 2001; Tschannen-Moran, Hoy, & Hoy, 1998).

Riggs and Enochs (1990) designed two science teaching efficacy instruments that were modelled on the Teacher Self Efficacy scales (TSES) produced by Gibson and Dembo (1984). The Science Teaching Efficacy Belief Instrument A (STEBI-A) was designed to measure the science teaching efficacy of inservice elementary teachers (Riggs & Enochs, 1990). The Science Teaching Efficacy Belief Instrument B (STEBI-B) was designed to measure the science teaching efficacy of preservice elementary teachers (Enochs & Riggs, 1990). These instruments are equivalent as the STEBI-B was designed by modifying the items from the original STEBI-A instrument to reflect the perspectives of preservice teachers. The STEBI-B was used in this research to track participants' personal and general science teaching efficacy over the course of four years.

Over the past 25 years science teaching efficacy has been established as an important measure in science education research (Deehan, 2016). Cross sectional research at the preservice level has linked STEBs with growth mindsets (Yilmaz-Tuzun & Topcu, 2008), student-centred teaching approaches (Gencer & Çakiroglu, 2007), non-linear, inquiry-based views on knowledge acquisition (Sunger, 2007) and higher levels of science content knowledge (Sarikaya, Çakiroglu, & Tekkaya, 2005). For inservice teachers, science teaching efficacy has been shown to correlate positively with the prevalence of constructivist teaching approaches (Lardy, 2011), including the promotion of student autonomy (Lucero, Valcke, & Schellens, 2013). A powerful argument can be made that, as a measure, science teaching efficacy relates to the current challenges within science education.

Within the STEBI-B literature there is some inconsistency in the use of the PSTE and STOE subscales. A major meta-analysis revealed that, in the majority of instances, the PSTE subscale has shown higher mean scores and growth rates than the STOE subscale (Deehan, 2016). The broad focus of the STOE subscale (i.e. the capacity of science teaching to influence teaching in a general sense) may be the main factor as extraneous variables (e.g. personal science experiences and histories) inherently influence responses and scores. It is common for the STOE subscale to dismissed for reasons such as 'teacher-centric' item structure (Bursal, 2010), low reliability (Andersen, Dragsted, Evans, & Sørensen, 2004; Velthuis, 2014), the external locus of control (McDonnough & Matkins, 2010) and preservice teachers' lack of relevant experience (Cannon & Scharmann, 1996). Such trends may be symptomatic of challenges of and contradictions within Bandura's conceptualisation of self-efficacy (Bandura, 1997). Theorists have argued that outcome expectancies are causally linked to, rather than clearly separated from, self-efficacy (Corcoran, 1995;

Eastman & Marzillier, 1984; Kirsh, 1995). Bandura has both accepted and rejected these claims in a contradictory fashion (Williams, 2010). It appears the lack of clarity may be related to researchers' preference for measuring the PSTE instead of the STOE (Deehan, 2016). At the university level, there is some evidence to support this choice as there is a strong-to-moderate correlation between the PSTE and STOE effect sizes of preservice teachers (Deehan, 2016). However, if the goal of university science programmes is to affect positive change in school systems, the choice to diminish or omit outcome expectancies is short-sighted as the PSTE and STOE growth correlation does not extend to inservice teachers (Pearson's $R = 0.121$). In the past decade, there have been renewed calls for researchers to address both subscales as it has been argued that teachers need to feel confidence that the task will have the desired impact and they will be able to perform the task to this standard (Bautista, 2011; Bautista & Boone, 2015; Williams, 2010). Such calls are not going unheard as meaningful consideration for the outcome expectancies of preservice teachers has become more widespread in recent years (Deehan, 2016). In line with views espoused by Bautista and Boone (2015), the research presented in this article addresses both the PSTE and STOE subscales of the STEBI-B instrument.

From theory to practice – sources of science teaching efficacy

With the constructs defined and the research contexts for STEBs established, it is imperative the authors reconcile the science courses with verified sources of science teaching efficacy. Bandura (1977, 1986) identified four key sources of efficacy; mastery experiences (ME), vicarious experiences (VE), verbal persuasion (VP) and emotional arousal (EA). He also argued that direct and practical MEs provide the most significant boosts to efficacy beliefs. D. H. Palmer (2006) extended the list of efficacy sources for preservice elementary science education to include: cognitive content mastery (CCM), cognitive pedagogical mastery (CPM) and simulated modelling (SM). Of these three additional efficacy sources, a sample of 190 preservice elementary teachers indicated that CPM was the main source of efficacy (D. H. Palmer, 2006).

The seven sources of efficacy are addressed in the educational design of the science courses. Table 2 connects each of the innovative practices with established efficacy sources to position this research within the theoretical framework. MEs and VEs are

Table 2. Pedagogical approaches and efficacy sources in SC108 and SC308.

Pedagogical Approach	SC108	SC308	Efficacy Sources
Alternative conception targeting	*		ME, VP, EA, CCM. CPM, SM
Cooperative Learning	*	*	ME, VP, VE, EA, CCM, CPM, SM
Constructivism	*	*	ME, VP, VE, EA, CCM, CPM, SM
ICT Instruction	*	*	CCM, CPM
Inquiry learning	*	*	ME, VP, CCM, CPM
In subject practical experience	*	*	ME, VP, VE, EA, CCM, CPM
Integration with other KLAs	*	*	CCM, CPM
Links to practical experience blocks			ME, VP, VE, EA, CCM, CPM
Mentoring	*		VP, VE, CPM, SM
Microteaching	*		ME, VP, VE, EA, CCM, CPM, SM
Nature of Science	*	*	ME, VP, CCM, CPM
Problem-based Learning	*		ME, VP, EA, CCM, CPM
Project-based Learning		*	ME, VP, CCM, CPM, SM
Real-world relevance	*	*	ME, VP, VE, EA, CCM, CPM
Student-centred Investigation	*	*	ME, VP, CCM, CPM

incorporated into the science programme via the deep inclusion of career specific tasks in a cooperative learning framework (e.g. microteaching experiences, mastery Astronomy learning, an in-school science education experience and science curriculum planning). VP is fostered through meaningful instruction in and use of cooperative learning strategies and reflective practices. EA is targeted at key stages throughout the programme to motivate preservice teachers. For example, participants are confronted by their own personal alternative conceptions at the beginning of the SC108 course. Positive arousal occurs when participants successfully design and implement curriculum in elementary settings. To promote preservice teachers' PCK (Nillson & Loughran, 2011) CCM and CPM are synthesised within the Interactive Educational Design Model of the science programme. In essence, participants are required to organically develop their PCK through progressively complex learning tasks reflecting the requirements of the teaching profession. SM is incorporated via early microteaching experiences, wherein participants rotate between the role of elementary teachers and students, and an extended project based learning scenario where participants and professors assume the roles teachers and principals respectively. The following paragraphs will expand on the educational design of the science courses.

Science courses

The first science course 'Science and Technology Studies I' (SC108) was positioned in the first semester of the first year of the degree. The second science course 'Science and Technology Curriculum Studies II' (SC308) was completed in the second semester of the second year of the degree. Table 2 identifies the innovative pedagogical approaches and efficacy sources in both courses. These innovative practices were initially expressed by Lawrance and Palmer (2003) and have been developed further by other researchers over the previous decade (Deehan, 2016). At a fundamental level, these innovative practices align with constructivist, student-centred approaches to learning rather than more traditional, teacher centred modes of content delivery. The authors define these courses as 'complex' due to the high number of interwoven innovative practices. Both SC108 (13) and SC308 (10) employ more than double the average number of innovative practices (4.6) reported within the STEBI-B literature (Deehan, 2016).

SC108 – Science and Technology Studies I

The first science subject used Astronomy science content as a driver for the development preservice teachers' PCK (Nillson & Loughran, 2011). The Astronomy content was mapped to the elementary science curriculum. A problem based learning environment was created by submitting the students to an Astronomy Diagnostic Test (ADT) (CAER, 1999) in the first week of class. The students were required to identify and target their own alternative scientific conceptions and to address these in cooperative, microteaching groups. The microteaching groups approach was inquiry based as students selected their own pedagogical approaches and reflected on their peer teaching. The success of the Astronomy teaching and learning was assessed by post-test administration of the ADT in the final week of the semester. Participants also developed their understandings of the nature of science by conducting investigations. The professor acted as a facilitator throughout the microteaching process.

SC308 – Science and Technology Curriculum Studies II

The knowledge and skills developed in SC108 were prerequisites for entry into the SC308 subject. SC308 moves beyond the Astronomy focus of SC108, as students were required to develop their PCK (Magnusson, Krajcik, & Borko, 2002; Van Driel & Berry, 2010) in order to teach syllabus content effectively. A hybrid constructivist pedagogical approach was used within an extended role play where the professor acted as the principal of the school (the tutorial class) and the preservice teachers assumed roles as teachers within the school. The preservice teachers work in cooperative learning groups throughout the semester (approximately four per group) to create a science unit of work for a different science content strand than the Earth and Space content covered in SC108. Each practical class had a different content focus and each cooperative learning group had a different stage focus (*F*, grades 1 and 2, grades 3 and 4, grades 5 and 6). The cooperative learning groups were required to navigate the syllabus, research science content and make pedagogical decisions to design a teachable science unit of work. The goal of this educational paradigm was to provide the preservice elementary teachers with the skills and knowledge necessary to research and adapt science concepts for the classroom. While these short summaries provide the reader with the core contextual information needed to interpret the research presented within this paper, delineating specific instructional approaches is particularly challenging as they are coalesced into an interactive educational design in which interactions are mediated through a PCK framework. A full description of the educational design and pedagogical inclusions of the SC108 and SC308 courses is the subject of another paper (McKinnon, Danaia, & Deehan, 2017).

Participants

The participants in this research were a cohort of 112 preservice elementary teachers enrolled in a Bachelor of Education (Elementary) degree at a regional Australian university. This degree would provide them with the knowledge, skills and qualifications necessary to teach in Australian elementary schools. The participants attended the university from 2010 through to 2013.

Due to the longitudinal nature of the research, attrition contributed to declines in response rates over time. On the first occasion of testing, 112 preservice teachers provided data. By the final occasion of testing 56 preservice teachers provided data. The age of the participants ranged from 18 to 55. The cohort was overwhelmingly female with 72% to 71% representation at the beginning and end of the degree respectively. A total of 12 members of the cohort participated in semi structured interviews.

The preservice teachers complete 32 subjects with 14 of these focusing on the six Australian elementary Key Learning Areas (KLAs) and elementary teaching pedagogies. To be awarded the degree the preservice elementary teachers had to complete the two science curriculum courses (SC108 and SC308).

Methodology and methods

Two concurrent nested mixed methods, Type II case studies (Burns, 2000; Creswell, 2013; Shadish, Cook, & Campbell, 2002; Yin, 2003) were conducted to explore the science experiences and STEBs of a cohort of preservice elementary teachers at an Australian university. A Type II case study involves the use of multiple methods at a single research site. The investigation used a repeated measures design with delayed testing periods (Shadish et al., 2002) to strengthen the case for covariance in the absence of an experimental design with a control group. Table 3 provides an outline of the treatment and data collection time frame. Quantitative STEBI-B data were collected weekly, alongside supplementary interview data, during the SC108 and SC308 courses. Additional data collection periods, denoted with DC in the table below, were positioned between delay periods and practical teaching experiences marked by the absence of a formal science education programme. This allowed for the participants' PSTE and STOE scores to be assessed within and beyond the courses.

Earlier iterations of both courses indicated that there were significant events during a semester which appeared to have an impact on the students. For example, in SC308, we had observed the impact of the in-school science teaching experience on them and wished to investigate further. Consequently, for this one cohort of preservice teachers, we administered the STEBI-B on an almost weekly basis in the two courses to generate 12 occasions in SC108 and 10 occasions in SC308 on which data were collected. Following these courses in the participants' third and fourth years of study, a further four administrations prior to and following their in-school professional experiences were conducted to generate a total of 26 occasions.

One could argue that such frequent administration of the instrument would contaminate the data through test–retest familiarity. We were careful to explain on each administration of the STEBI-B that students should read and carefully respond to each item in light of their *current* feelings. Most students appeared to treat each completion of the instrument in a conscientious fashion while some made comment in their final evaluations that they never wished to see it again. Given the extensive reliability analyses, we believe that the changes in both scales are real and do not reflect any major threat to validity caused by test familiarity. The following subsections will unpack each of the data collection methods in greater depth.

The Science Teaching Efficacy Belief Instrument (B)

The Science Teaching Efficacy Belief Instrument B (STEBI-B) was used to measure the science teaching efficacy of preservice elementary teachers (Enochs & Riggs, 1990).

Table 3. Treatment and data collection.

2010		2011		2012		2013	
First Semester	Second Semester	First Semester	Second Semester	First Semester	Second Semester	First Semester	Second Semester
Entered teaching degree Undertook SC108 (DC)	Delay period Practical teaching experience	Delay period	Undertook SC308 (DC) Practical teaching experience	Delay period Follow-up (DC)	Delay period Practical teaching experience	Delay period Follow-up (DC)	Practical teaching experience Follow-up (DC) Completed teaching degree

The instrument requires respondents to rate their level of agreement with statements on a 5-point Likert scale (Burns, 2000), ranging from 'strongly disagree' to 'strongly agree'. The statements produce measurements on two subscales. The Science Teaching Outcome Expectancy (STOE) belief scale measures the participants' broad views of science teaching related to why pupils perform as they do. The Personal Science Teaching Efficacy (PSTE) scale measures the participants' beliefs about their own ability to teach science effectively. The seminal authors reported Cronbach's Alpha reliability coefficients of 0.90 for the PSTE subscale and 0.76 for the STOE subscale (Enochs & Riggs, 1990).

The comparatively low reliability and external locus of the STOE subscale have emerged as contentious issues within the broader body of STEBI-B literature. The STOE subscale has been reported as having lower reliability than the PSTE subscale in a variety of contexts (e.g. Aydin & Boz, 2010; Bleicher, 2006; Riggs & Enochs, 1990; Velthuis, Fisser, & Pieters, 2014). McDonnough and Matkins (2010) expressed doubt in the reliability of the STOE subscale due to the external locus of control. Indeed, others (e.g. Bursal, 2008; Hechter, 2010) believe that the sheer volume of potential influencing factors make the STOE subscale conceptually unclear. Mulholland et al. (2004) believe that the items comprising the STOE subscale reflect an outdated, teacher centred mode of science teaching. They believe that inconsistency in participants' responses can be partially attributed to an inability to relate to the items to their experiences in modern teacher preparation programmes. The STOE scale is often dismissed (e.g. Andersen et al., 2004; Cannon & Scharmann, 1996) or merged with the PSTE subscale (e.g. Slater, Slater, & Shaner, 2008). Merging subscales is a particularly inappropriate practice as it both denies the separate nature of constructs and artificially inflates alpha scores through a greater number of items (Tavakol & Dennick, 2011). The authors believe that broader contextual efficacy is necessary for an individual to persevere in their completion of a task in cases of adversity. Thus, the reliability of the STEBI-B instrument was measured repeatedly in the four-year data collection period.

Although still acceptable, the reliability of the subscales appeared to be an issue early in the current research as the PSTE (Cronbach's $\alpha = 0.73$) and the STOE (Cronbach's $\alpha = 0.69$) reliabilities were well below those reported by seminal authors. Later investigations showed that the reliability of the scales improved over the course of the four-year data collection period. On the final occasion of testing, just prior to the cohort's graduation, the PSTE and STOE subscales showed Cronbach's alpha reliabilities of 0.88 and 0.87 respectively. Such equality in subscale reliability is seldom reported in the STEBI-B literature. The implications of this information will be analysed further in the discussion section of this paper.

Semi structured interviews

The semi structured interviews within the current research explored the science experiences, perceptions, beliefs and opinions of selected preservice elementary teachers in relation to their participation in the science curriculum courses (SC108/ SC308) and their emerging capacities as science teachers. A combination of convenience and snowball sampling procedures were used to gain access to willing participants. Repeated semi structured interviews were conducted with 12 participants, during the delivery of SC108, SC308 and on the first delayed testing occasion.

Data analyses

For the STEBI-B data, Multivariate Analysis of Variance (MANOVA) with repeated measures on the occasion of testing were computed to determine if the PSTE and STOE scores of participants changed significantly over the four-year period (Coleman & Pilford, 2008). Cohen's *d* was computed to determine the effect sizes of any STEB changes occurring over time. In addition, SPSS statistical software was used to analyse and produce descriptive statistics, tests for the homogeneity of variance and tests for the equality of the covariance matrices to ensure that the distributions of the DVs met the mathematical assumptions of MANOVA. In accordance with the Multivariate Central Limit Theorem, there are at least 20 elements for each variable combination within the analyses. This means the multivariate normality assumption can be held. Analysis via Hotelling's T^2 tests confirmed the normal distribution of the dataset.

All qualitative data were transcribed and manually analysed by the lead author. A thorough process of listening to audio files, manual transcription and multiple readings of the transcripts was undertaken to identify key words, concepts and ideas (Hsieh & Shannon, 2005). Both core themes, directly linked to the research aims, and emergent themes, unanticipated but relevant findings, were considered during this process (O'Toole & Beckett, 2013). QSR NVIVO 10 software housed the coded themes to allow for the relationships between the themes to be interrogated objectively (Bryman, 2016). The next step was to recode the data for consistency and clarity without altering the meaning of the text. To avoid bias and researcher error, these analyses were supplemented with a computer analysis using Leximancer software. The Leximancer software was used to conduct broad, syntactical analyses of the text to identify themes and to explore thematic relationships. This semantic analysis served as a useful complement to the lead researcher's personal identification of themes and was a good test for the validity of interpretation. A rigorous interrater reliability process was undertaken where the coding of the lead researcher was cross checked for consistency by the second and third researchers. Students' emotions, science programme responses and in-school practical teaching experiences were identified by the researchers as key categories from which connections to literature, answers to questions and emergent sub-themes could be derived.

Findings

This section has been organised to address each of the separate aims and questions before presenting a holistic overview of the findings.

The first aim is to assess the relationship between preservice elementary teachers' participation in a complex, integrated science programme with two complementary courses (SC108 and SC308) and their STEBs. The second aim is to determine if any changes to participants' STEBs that occur within the science programme remain durable for up to two years in the absence of treatment. The third aim is to inductively explore the participants' attitudes towards and perceptions of science teaching as they progressed through the science courses and beyond.

Question 1 – What is the relationship between preservice elementary teachers' participation in a complex, integrated science programme with two complementary courses (SC108 and SC308) and their STEBs?

Answer

STEBI-B data show a positive covariant relationship between participation in the SC108 course and preservice elementary teachers' STEBs. Table 4 presents the results for the MANOVA conducted on the pre to post occasion STEB data collected in SC108. There is a significant main effect due to the occasion of testing ($F(1{,}107) = 52.94$, $p < .0001$). This indicates that the PSTE and STOE scores of the participants showed a statistically significant increase as they undertook the SC108 subject. There is also a significant main effect due to the variables, PSTE versus STOE ($F(1{,}107) = 42.85$, $p < .0001$). A possible interpretation may be that the preservice teachers did not yet feel confident that they could personally fulfil their own broader expectations of science teaching to improve student outcomes. There was no significant interaction effect due to occasion with variable ($F(1{,}107) = 0.75$, $p = .389$).

Positive trends were also evident in the science teaching efficacy of the cohort throughout the SC108 semester. Figure 1 shows the mean PSTE and STOE scores of the participants as they progressed through the SC108 course. There was a consistent, if uneven, increase in both subscales throughout the course. Of note is the finding that the STOE means were higher than the PSTE means on all occasions of testing. This is an anomaly as the PSTE is almost always higher than the STOE subscale in the STEBI-B literature (Deehan, 2016). The PSTE showed a moderate effect size increase (Cohen's $d = 0.41$). The STOE subscale showed a large effect size increase (Cohen's $d = 0.79$).

There was highly significant growth in the cohort's science teaching efficacy scores upon completion of the SC308 course. Table 5 shows the statistical output for the MANOVA conducted on the SC308 STEB data. There is a significant main effect due to the occasion of testing ($F(1{,}46) = 34.08$, $p < .0001$). There is also a significant main effect due to the variables, PSTE versus STOE ($F(1{,}46) = 22.75$, $p < .0001$). That is to say, the gap between their outcome expectancies and personal STEBs remained present throughout the SC308 course. There was no significant interaction effect due to occasion with variable ($F(1{,}46) = 2.45$, $p = .125$).

The trend of growth in the STEBs subscales continued as the cohort worked through SC308. Figure 2 shows the mean PSTE and STOE scores of the participants for each occasion of testing in the SC308 course. There was an unprecedented increase of the PSTE scale during the eighth week of the semester. The increase coincided with the preservice teachers' participation in a science teaching day where they taught some of the content of their units of work to groups of students drawn from a number of local schools. Overall, the PSTE subscale showed a moderate effect size increase

Table 4. MANOVA of STEB data collection during the SC108 course.

Variable	*SS*	*df*	*MS*	*F*	*p*
Occasion	702.78	1	702.78	52.94	<.0001
Error(Occasion)	1420.47	107	13.28		
Variable	770.67	1	770.67	42.85	<.0001
Error(Variable)	1924.58	107	17.99		
Occasion * Variable	7.00	1	7.00	.75	.389
Error(Occasion * Variable)	1002.25	107	9.37		

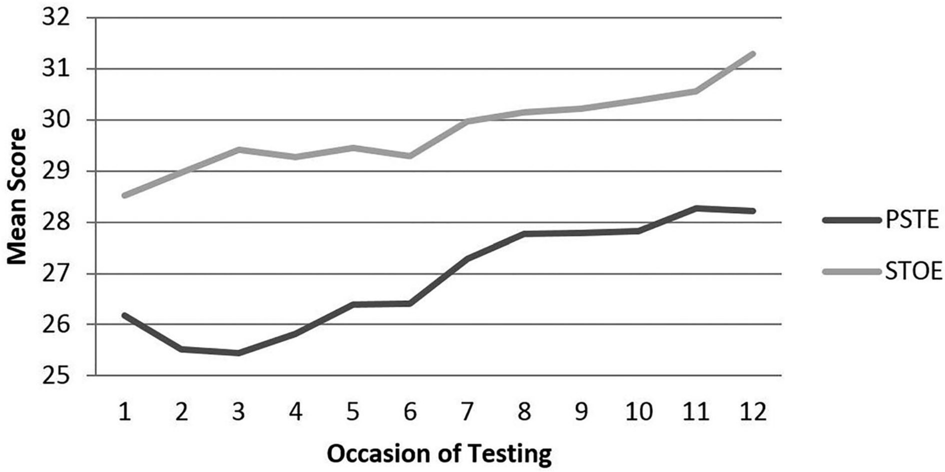

Figure 1. SC108: Mean STEB scores by week.

(Cohen's $d = 0.65$). The STOE subscale underwent a similar increase (Cohen's $d = 0.69$) during the SC308 subject.

Question 2 – If participants' STEBs change during the science programme (SC108 and SC308), do these changes remain durable for up to two years in the absence of treatment? Answer

The preservice teachers' STEBs remained stable during the first year with no formal science intervention. Table 6 shows the descriptive statistics for the STEB scales collected at the end of SC308, prior to and following a practical teaching placement (delay period one and delay period two). Clearly, there was very little change in their STEBs despite no formal science education for a year. A MANOVA with repeated measures showed that there was no significant difference in participants' STEBs on the occasion of testing ($F(1,29) = 1.98$, $p = .167$) during this period.

Not only did the preservice teachers' STEBs remain durable, the mean STEB scores increased despite the absence of any science intervention for the intervening two-year period. Table 7 shows the descriptive statistics for the STEB data collected at the end of 2012 (delay period two), pre internship 2013 (delay period three) and post internship 2013 (delay period four). The preservice teachers' personal science teaching efficacy showed a small increase during their final year of study (Cohen's $d = 0.26$). The cohort's STOE beliefs showed a greater increase (Cohen's $d = 0.41$).

Table 5. MANOVA of STEB data collected during the SC308 course.

Variable	*SS*	df	*MS*	*F*	*p*
Occasion	550.94	1	550.93	34.08	<.0001
Error(Occasion)	743.62	46	16.12		
STEB	1973.45	1	1973.45	22.75	<.0001
Error(STEB)	3990.85	46	86.76		
Occasion * STEB	24.51	1	24.51	2.45	.125
Error(Occasion * STEB)	460.91	46	10.02		

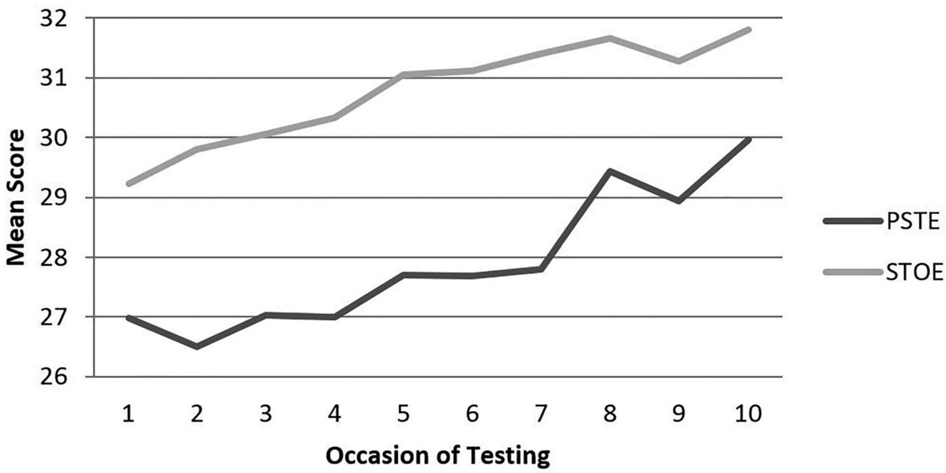

Figure 2. SC308: Mean STEB scores by week.

A MANOVA with repeated measure was computed on the STEB data collected during the preservice teachers' final year of tertiary studies. Table 8 presents the results from the MANOVA for this period. There was a significant effect on participants' STEBs due to occasion of testing ($F(1.68,53.83) = 8.165$, $p = .001$). This indicates that the growth in both the PSTE and STOE scores of the participants was statistically significant. An intriguing finding was that there was no significant difference between the PSTE and STOE subscales ($F(1,32) = 1.55$, $p = .222$). This was the first time during the four-year data collection period where the preservice teachers' PSTE beliefs were not significantly lower than their STOE beliefs. One interpretation of these results may be that their tertiary education had helped them to feel that they can meet their high standards for science teaching in general.

Table 6. Descriptive statistics for the end of SC308, delay period 1 and delay period 2.

Occasion	PSTE Occasion	*M*	*SD*	STOE Occasion	*M*	*SD*	*n*
END SC308	PERS24	29.96	4.61	GEN24	31.8	3.57	70
	2nd Year Practical Teaching Experience and Summer Break (6 months)						
Delay Period 1	PERS25	30.03	3.79	GEN25	31.7	2.72	39
	2012 – 1st Semester and 3rd Year Practical Teaching Experience (6 months)						
Delay Period 2	PERS26	30.90	4.00	GEN26	31.1	3.04	42

Table 7. Descriptive statistics for delay period 2, delay period 3 and delay period 4.

Occasion	PSTE Occasion	*M*	*SD*	STOE Occasion	*M*	*SD*	*n*
Second Semester 2012	PERS26	30.90	4.00	GEN26	31.1	3.04	42
	2012 Summer Break and 1st Semester 2013 (6 months)						
Pre Internship 2013	PERS27	31.13	4.67	GEN27	33.06	3.01	48
	2013 Teaching Internship (3 months)						
Post Internship 2013	PERS28	32.02	4.60	GEN28	32.57	4.03	55

Table 8. MANOVA output for the final year of study.

Variable	*SS*	*df*	*MS*	*F*	*p*
Occasion	152.19	1.68	90.47	8.17	.001
Error(Occasion)	596.48	53.83	11.08		
STEB	23.35	1	23.35	1.55	.222
Error(STEB)	481.65	32	15.05		
Occasion * STEB	21.28	2	10.64	3.39	.040
Error(Occasion * Variable)	200.72	64	3.14		

Question 3 – What were the participants' attitudes towards and perceptions of science teaching as they progressed through the science subjects and beyond?
Answer

The semi structured interview data reveal that by the end of SC108 the preservice teachers seemed to be more confident in their ability to teach science. They attributed their changed attitudes and improved STEBs to their experiences within the SC108 course. Cooperative learning, curriculum building and microteaching were identified as influential components. For example:

Malcolm: The big moment where I realised that I'd actually be confident teaching, like last night when I realised exactly how much I'd sort of done in the course. I read through my assignment and went, it's all for a reason. The curriculum building, the readings we did, the extra homework, all the cooperative learning strategies that we did in class. It all sort of came together in my head last night. I sort of had an epiphany thing, where I realised how useful this course is, to us going teaching. It's built my confidence up something fierce.

Aaron: I think it is the micro teaching part. We were able to put theory into practice, if you get what I mean. We learned about teaching strategies, we read and learned content knowledge and then we were able to put this into practice by trying and teaching the stuff. It was like I am going to try this, and I did, then it was like, hey that was brilliant or hey, no that sucked, what could I do to improve that?

In the 12 month period between the end of SC108 and the beginning of SC308 the STEBs of the preservice teachers declined. The PSTE declined by a small Cohen's d effect size of 0.12. More troubling was the moderate STOE decline (Cohen's $d = -0.56$). A MANOVA with repeated measures revealed that the decline in both the PSTE and STOE by occasion were significant ($F(1,61) = 9.71$, $p = .003$). The decline in the STOE may be related the preservice teachers' observations on their first practical teaching experiences. The avoidance of science in schools was a prominent theme in the semi structured interviews. For example, responding to the question 'Did you see any science being taught at your school?' the following quotes were elicited:

Edward: I'm pretty sure that there was no science taught. I'm pretty sure that it was embedded in there, but I didn't see it there.

Bec: No! Not one bit of science at all.

Daisy: No, no (science). None whatsoever.

Connie: I taught a couple of lessons which were called 'Science', but I wouldn't call them Science. They'd just done happy healthy Harold. They were lessons that were sort of half health, half science. They were about food groups and what goes into certain food groups. They called them 'Science lessons', but they didn't have a lot to do with the science curriculum.

Fiona: (Um) I honestly don't think so. I think it happened when we had release time, like relief from face to face [teaching], so I had to go with my associate then. So I never actually saw science being taught.

The learning experiences within the SC308 had a profound effect on the participants. They developed practical inquiry skills that would enable them to plan and teach quality science lessons. One interviewee discussed her development during the SC308 course:

Daisy: The fact that I have completed a whole unit of work makes me realise that I can actually teach science. The kids can get hands on, the experiments, they can see the answer for themselves while doing something hands on. It's really great and I hope when I go on prac that one of the lessons that I will teach will be science.

Whether or not the interviewees were aware, deeper analysis reveals that they may have been socialised into a culture of science marginalisation while on extended practical teaching placements. In one short response, Bec cited a plethora of reasons for dismissing science despite addressing it in her own curriculum:

> Just all the setting up. Planning was a bit more difficult too and having to get your resources for it. I had to get everything. That's also money from me, although I'm sure the school would have been happy to [pay]. With English and Maths you've got to cover so much and you want to get it ticked off before you start science. It is time consuming. Especially when you could just do an art lesson, like everyone does.

Conversely, another interviewee was highly critical of the science she observed during the placement. She chose to avoid common marginalisation processes as she made informed pedagogical and content decisions to deliver her science lessons. Fiona said:

> The content I taught was not as scientific as I had hoped. My associate teacher was very restrictive about what was to be taught. She just gave us a black and white printed off booklet and told us to teach it as she would which is basically colouring in the pictures in the book. My paired prac partner and I changed the book around and only kept some pages to group the content together and instead conducted experiments and SmartBoard activities that were meaningful rather than simply giving them sheets to colour in. The content we were originally given did not fit the Kindergarten syllabus at all so our activities steered the content so that it could actually satisfy the outcomes.

Summary of findings

These data suggest that participation within the two science courses (SC108 and SC308) covaried with statistically significant increases in both the PSTE and STOE scores for the cohort of preservice elementary teachers. These improved STEBs remained durable after, and even increased in the absence of, the science subjects.

Figure 3 shows the STEB progression over the four-year period, based on the mean scores for data provided at each occasion. The solid vertical lines represent the beginning of an academic year and the dotted lines represent the end of an academic year. Throughout the first two years of the degree, where the science courses were situated in the course structure, the outcome expectancies of the preservice teachers were consistently higher than their personal STEBs. This trend appears logical as during these early stages of their tertiary education, they did not possess the experience, knowledge and skills of qualified teachers nor a real context within which they could observe the effects of teacher

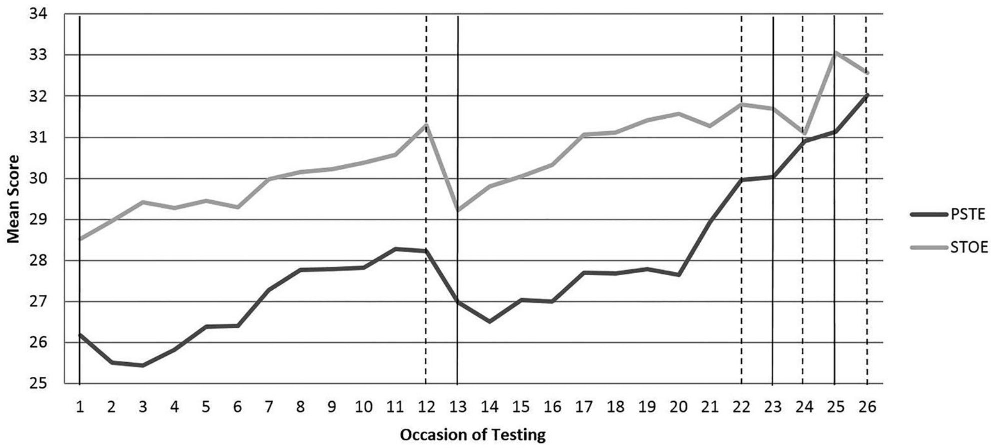

Figure 3. Progression of STEB scores over the four-year period.

effort. It should be noted that this difference between the STOE and PSTE scores is relatively unique in the literature as many researchers consistently report both higher scores and higher growth on the PSTE subscale (e.g. Ford, Fifield, Madsen, & Qian, 2012; Logerwell, 2009; D. Palmer, 2006). Another noteworthy trend was the continued increase in science teaching efficacy belief scores after the completion of the science courses. The gap between the PSTE and STOE scores also decreased, indicating that the preservice teachers now felt confident that they could deliver quality science teaching on a par with the broader science teaching profession.

To gain a holistic overview of the combined influence of both the science courses (SC108 and SC308) and the other aspects of the teaching degree on the cohort's STEBs a MANOVA was computed on the STEB data collected at the beginning of the degree (SC108 entry) and at the end of the degree (post internship). 45 preservice teachers provided valid STEBI-B data on both occasions. Table 9 shows that there was a significant effect due to the STEB scales ($F(1,44) = 6.88$, $p = .012$). The difference between the PSTE and STOE scales was more pronounced upon entry into the degree and this is reflected in the significant effect due to STEB by Occasion ($F = (1,44) = 6.88$, $p = 0.12$). There was a highly significant effect due to Occasion ($F(1,44) = 82.58$, $p < .0001$). The PSTE and STOE subscales both showed large-to-very-large Cohen's d effect sizes of 1.36 and 1.11 respectively.

Table 9. MANOVA of STEB data collected on the first and final occasions of testing over the four-year period.

Variable	*SS*	*df*	*MS*	*F*	*p*
Occasion	1120.01	1	1120.01	82.58	<.0001
Error(Occasion)	596.74	44	12.56		
STEB	81.34	1	81.34	6.88	.012
Error(STEB)	530.41	44	11.83		
Occasion * STEB	61.25	1	61.25	6.88	.012
Error(Occasion * STEB)	391.5	44	8.90		

Discussion

The research presented in this paper shows that participation in two complementary science courses, comprising multiple innovative practices, covaries with improved STEBs and which are durable in the absence of treatment. Across both science courses, participants displayed large effect size increases on both the PSTE (Cohen's $d = 0.75$) and STOE (Cohen's $d = 1.08$) subscales. Assessing both science courses as a part of a broader science programme may be the most suitable approach due to the statistically significant decline in STEB scores that occurred in the delay period between the first (SC108) and second (SC308) course. After the cohort had completed SC308, their STEBs remained durable for two years without any formal science intervention. In fact, a pre- and post-course MANOVA showed large effect size increases in the PSTE (Cohen's $d = 1.36$) and STOE (Cohen's $d = 1.11$). It appears the preservice teachers maintained their beliefs about the efficacy of science teaching whilst developing their personal STEBs through continued teaching experience and study. Any interpretation must be treated with circumspection as extraneous factors such as maturation, other courses within the degree and general practical teaching experiences are all likely to contribute to preservice teachers' efficacy beliefs.

The weekly collection of data within the science courses allowed for some insights into how different pedagogical innovations affected the STEBs of the preservice teachers. The embedded practical science teaching experience in the second science course (SC308) covaried with the largest rise in PSTE scores within the data set. In a single week, the participants displayed a small, but significant, increase (Cohen's $d = 0.25$) in their personal science teaching efficacy scores. Many researchers report on the benefits of including practical teaching experiences within tertiary science courses (e.g. Bautista, 2011; Lewthwaite, Murray, & Hechter, 2012; D. Palmer, 2006; Velthuis et al., 2014). The findings in this research represent a substantial advancement in this area of research as the weekly STEBI-B administrations allowed for the practical teaching event to be isolated from other components of the SC308 course. This research provides reasonable evidence for the benefits of the inclusion of practical science teaching experiences in tertiary science education subjects. Nonetheless, more research is needed to strengthen such an argument. The benefits of practical science teaching experience could also open avenues for partnerships between universities and schools where such opportunities could serve to counteract any negative science socialisation that appears to occur for both preservice and early career teachers in schools. Some research has been conducted on linking tertiary science courses to extended professional experience placement (e.g. Bautista, 2011; Leonard, Barnes-Johnson, Dantley, & Kimber, 2011; McDonnough & Matkins, 2010), but further research needs to be conducted in this potentially rich domain.

The interview data collected within this research serves to highlight the negative science socialisation that appears to occur within some elementary schools. Such themes certainly fit with the negative trends that emerge within the wider literature base. Elementary science is often marginalised (e.g. Appleton, 2003; Appleton & Kindt, 2002; Angus et al., 2004) and distorted (e.g. Goodrum & Rennie, 2007) by teachers who lack confidence in their capacity to teach science effectively (e.g. Palmer, 2011; Tytler et al., 1999). All the interviewees noted the absence of science, but their interpretations of the science avoidance differed in two distinct ways. Some of the interviewees were highly critical of science marginalisation and chose to eschew the power structure that existed between themselves and their associate teachers by teaching science in a student-centred way. Others seemed compelled to accompany their answers with rationalisations on behalf of their associate teachers. Issues such as time, resourcing and the crowded curriculum became key themes for the preservice teachers despite their own lack of experience within the teaching profession. Limited research exists that directly reports on science cultures at the elementary level, but the dissonance which exists in the research hints at negative science cultures. For example, teachers who participated in a national study into the state of science education reported using student-centred approaches, yet the elementary students reported transmissive, teacher centred pedagogies (Goodrum et al., 2001). In the same report, many elementary teachers cited the lack of professional development opportunities in science teaching. Yet, in a follow-up study, many elementary teachers displayed a lack of interest in engaging with professional development opportunities (Goodrum & Rennie, 2007). Given that any science gains made at the tertiary level will inevitably interact with the school cultures into which participants enter, research needs to be conducted to study science socialisation at the elementary level.

Throughout the four-year research period, the STOE scores of the participants were consistently higher than their PSTE scores. It is logical that inexperienced preservice teachers would feel less efficacious in their own science teaching capacity than the capacity of science teaching to assist student learning in a general sense. Nonetheless, the overwhelming majority of the STEBI-B literature reports higher scores and growth on the PSTE subscale (Deehan, 2016; Ford et al., 2012; Logerwell, 2009; D. Palmer, 2006). It is impossible to attribute this trend to any single factor within the science programme or indeed within the four-year data collection period. It could be posited that the alternative conception targeting that occurred early in SC108 made the preservice teachers more acutely aware of the limitations of their science content knowledge. Additionally, the student centred, microteaching environment showed them that teaching science requires a wide array of professional skills which they, as first year preservice teachers, did not yet possess. Such an interpretation is supported by the reduction in the gap between the subscales that occurred as the cohort progressed through their degree. By the time the group was ready to graduate, there was no longer a statistically significant gap between their PSTE and STOE scores. This hints at a powerful narrative, wherein the cohort believed that their tertiary experiences had prepared them to meet their own expectations of elementary science teaching. Nonetheless, follow up is research is needed to determine the validity of this interpretation.

The validity and reliability of the STOE subscale remains a point of contention within the STEBI-B literature (e.g. McDonnough & Matkins, 2010; Mulholland et al., 2004). As noted earlier, the STOE is often marred by low reliability scores (e.g. Aydin & Boz, 2010; Velthuis et al., 2014) and subsequently removed from reporting. The current research afforded an opportunity to explore the reliability of the STOE subscale in a longitudinal manner. As the targeted cohort of preservice elementary teachers progressed through their degrees, the reliability of their responses to the STOE items increased. At the beginning of the first year, the STOE was reported to have a Cronbach's alpha of 0.687, by the end of the second year this had increased to 0.798. Most notably, despite no formal science intervention during the interim period, the STOE Cronbach's alpha had increased to 0.87 and at the end of the fourth year, which was very close the 0.88 reliability score of the PSTE subscale. These reliability scores represent one of the only times in the literature where both the PSTE and STOE subscales display equal reliability. This reaffirms the importance of practical teaching experiences both within and beyond tertiary science programsme. There is some evidence here to suggest that the STOE reliability increases naturally over the course of a tertiary teacher education programme, as preservice teachers' develop deeper understandings of the teaching profession. While the aforementioned STOE issues are valid, they are not infallible nor do they remove the imperative to improve preservice elementary teachers' beliefs about the capacity of science teaching to guide students to desired learning outcomes. Simply put, educators must continue to improve preservice teachers' science teaching outcome expectancies.

There are three key limitations to the research presented within this paper. First, the reliance on a quasi-experimental, longitudinal design does not allow for a causal relationship between the science courses (SC108 and SC308) and participants' STEBs. The absence of an experimental design prevents extraneous variables (educational background, personal experiences, etc.) from being dismissed or even considered within the relationship. Second, the longitudinal design increases the likelihood of survey fatigue (Shadish et al., 2002) negatively influencing the accuracy and consistency of participants' responses to the STEBI-B. Third, despite the research focus on the durability of STEB gains in this paper, the context does not extend into the participants' transition from preservice to inservice status. So while the benefits of improved STEBs can be inferred from existing literature, there is yet no evidence of improved science teaching and learning within schools.

There are several implications for further research that have been revealed by the research outlined in this paper. Follow-up research is needed to explore the STEBs and science teaching practices of the preservice elementary teaching cohorts as they begin their teaching careers. A disconnect exists between tertiary and elementary school contexts as research seldom bridges this gap (McKinnon & Lamberts, 2014). It is necessary to extend research beyond the tertiary context if the goal of improving students' scientific literacy is to be addressed. Research also needs to be conducted explicitly on the negative science cultures that appear to exist in elementary schools. Such research would provide insights into the types of interventions required to overcome the issues of science avoidance and marginalisation. From a methodological standpoint, more research needs to adopt multiple cohort designs, longitudinal approaches and good quasi-experimental designs. This would improve arguments for covariance amongst variables and potentially establish causal links between science interventions and outcomes.

Conclusion

The evidence presented within the paper shows that the science programme (SC108 and SC308) represent a viable model for improving the STEBs of preservice elementary teachers. The participants' significant STEB growth remained durable for two years without any formal science intervention. The high mean scores, strong growth rates and subscale equity suggest that the preservice teachers held strong beliefs that their science teaching could overcome external barriers to assist students to meet outcomes. Indeed, an interpretation of both the qualitative and quantitative data may be that their later practical experiences and continued university learning consolidated, rather than threatened, their high outcome expectancies. Consistently high science teaching outcome expectancies will be crucial if these individuals are to display professional resilience in response to the challenges facing contemporary science education. Most importantly, they completed their degrees without significant disparities between their PSTE and STOE scores. That is to say, the preservice teachers felt capable of fulfilling their own high expectations of science teaching in general. Clearly, the SC108 and SC308 courses are strong alternatives to more traditional approaches to preservice science education as they imbue the participants with resilient attitudes and durable STEBs.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

James Deehan http://orcid.org/0000-0001-5825-9734
Lena Danaia http://orcid.org/0000-0002-1924-3905
David H. McKinnon http://orcid.org/0000-0001-9549-9921

References

Akerson, V. L., Pongsanon, K., Rogers, M. A. P., Carter, I., & Galindo, E. (2017). Exploring the use of lesson study to develop elementary preservice teachers' pedagogical content knowledge for teaching nature of science. *International Journal of Science and Mathematics Education*, *15*(2), 293–312.

Andersen, A. M., Dragsted, S., Evans, R. H., & Sørensen, H. (2004). The relationship between changes in teachers' self-efficacy beliefs and the science teaching environment of Danish first-year elementary teachers. *Journal of Science Teacher Education*, *15*(1), 25–38.

Appleton, K. (1992). Discipline knowledge and confidence to teach science: Self-perceptions of primary teacher education students. *Research in Science Education*, *22*, 11–19.

Appleton, K. (2002). Science activities that work: Perceptions of primary school teachers. *Research in Science Education*, *32*(3), 393–410.

Appleton, K. (2003). How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in Science Education*, *33*(1), 1–25.

Appleton, K., & Kindt, I. (2002). Beginning elementary teachers' development as teachers of science. *Journal of Science Teacher Education*, *13*(1), 43–61.

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

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

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: Freeman.

Bautista, N. U. (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.

Bautista, N. U., & Boone, W. J. (2015). Exploring the impact of teachME™ lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education*, *26*(3), 237–262.

Bleicher, R. E. (2006). Nurturing confidence in preservice elementary science teachers. *Journal of Science Teacher Education*, *17*(2), 165–187.

Bleicher, R. E., & Lindgren, J. (2005). Success in science learning and preservice science teaching self-efficacy. *Journal of Science Teacher Education*, *16*(3), 205–225.
Bryman, A. (2016). *Social research methods* (5th ed.). New York, NY: Oxford University Press.
Burns, R. (2000). *Introduction to research methods* (3rd ed.). Melbourne: Longman Publishers.
Bursal, M. (2008). Changes in Turkish pre-service elementary teachers' personal science teaching efficacy beliefs and science anxieties during a science method course. *Journal of Turkish Science Education*, *5*(1), 99–112.
Bursal, M. (2010). Turkish preservice elementary teachers' self-efficacy beliefs regarding mathematics and science teaching. *International Journal of Science and Mathematics Education*, *8*(4), 649–666.
Cannon, J. R., & Scharmann, L. C. (1996). Influence of a cooperative early field experience on preservice elementary teachers' science self-efficacy. *Science Education*, *80*(4), 419–436.
Coladarci, T. (1992). Teachers' sense of efficacy and commitment to teaching. *The Journal of Experimental Education*, *60*(4), 323–337.
Coleman, A., & Pilford, B. (2008). *A crash course in SPSS for windows* (4th ed.). West Sussex: Blackwell Publishing.
Collaboration for Astronomy Education Research (CAER). (1999, June 21). *Astronomy diagnostic test (ADT) version 2.0*. Retrieved from Association for Astronomy Education: http://solar.physics.montana.edu/aae/adt/ADTv2.0.PDF
Cooper, G., Kenny, J., & Fraser, S. (2012). Influencing intended teaching practice: Exploring preservice teachers' perceptions of science teaching resources. *International Journal of Science Education*, *34*(12), 1883–1908.
Corcoran, K. J. (1995). Understanding cognition, choice, and behavior. *Journal of Behavior Therapy and Experimental Psychiatry*, *26*, 201–207.
Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. London: Sage Publications.
Deehan, J. (2016). *The science teaching efficacy belief instruments (A and B): A comprehensive review of methods and findings from 25 years of science education research*. Cham,Switzerland: Springer International Publishing.
de Laat, J., & Watters, J. J. (1995). Science teaching self-efficacy in a primary school: A case study. *Research in Science Education*, *25*(4), 453–464.
DeWitt, J., Archer, L., & Osborne, J. (2014). Science-related aspirations across the primary–secondary divide: Evidence from two surveys in England. *International Journal of Science Education*, *36*(10), 1609–1629.
Duran, E., Ballone-Duran, L., Haney, J., & Beltyukova, S. (2009). The impact of a professional development program integrating informal science education on early childhood teachers' self-efficacy and beliefs about inquiry-based science teaching. *Journal of Elementary Science Education*, *21*(4), 53–70.
Eastman, C., & Marzillier, J. S. (1984). Theoretical and methodological difficulties in Bandura's self-efficacy theory. *Cognitive Therapy and Research*, *8*, 213–229.
Enochs, L. G., & 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.
Ford, D. J., Fifield, S., Madsen, J., & Qian, X. (2012). The science semester: Cross-disciplinary inquiry for prospective elementary teachers. *Journal of Science Teacher Education*, *24*(6), 1049–1072.
Gencer, A. S., & Çakiroglu, 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.
Gibson, S., & 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., & Lucas, K. B. (1995). Changes in preservice elementary teachers' sense of efficacy in teaching science. *School Science and Mathematics*, *95*(8), 394–400.

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2000). Collective teacher efficacy: Its meaning, measure, and impact on student achievement. *American Educational Research Journal*, *37*(2), 479–507.
Gonski, D. (2011). *Review of funding for schooling*. Canberra: Commonwealth of Australia.
Goodrum, D., Hackling, M., & Rennie, L. (2001). *The status and quality of teaching and learning of science in Australian schools*. Canberra: Department of Education, Training and Youth Affairs.
Goodrum, D., & Rennie, L. (2007). *Australian school science education – national action plan 2008–2012* (Vol. 1). Canberra: Commonwealth of Australia.
Gordon, C., & Debus, R. (2002). Developing deep learning approaches and personal teaching efficacy within a preservice teacher education context. *British Journal of Educational Psychology*, *72* (1), 483–511.
Gough, A., Marshall, A., Matthews, R., Milne, G., Tytler, R., & White, G. (1998). *Science baseline survey research report*. Melbourne: Deakin University.
Guskey, T. R., & Passaro, P. D. (1994). Teacher efficacy: A study of construct dimensions. *American Educational Research Journal*, *31*, 627–643.
Harris, K. L., & Farrell, K. (2007). The science shortfall: An analysis of the shortage of suitably qualified science teachers in Australian schools and the policy implications for universities. *Journal of Higher Education Policy and Management*, *29*(2), 159–171.
Hechter, R. (2010). Changes in preservice elementary teachers' personal science teaching efficacy and science teaching outcome expectancies: The influence of context. *Journal of Science Teacher Education*, *22*(1), 187–202.
Howitt, C. (2007). Pre-service elementary teachers' perceptions of factors in an holistic methods course influencing their confidence in teaching science. *Research in Science Education*, *37*(1), 41–58.
Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288.
Jarrett, O. (1999). Science interest and confidence among preservice elementary teachers. *Journal of Elementary Science Education*, *1*(11), 47–57.
Juriševič, M., Glažar, S. A., Pučko, C. R., & Devetak, I. (2008). Intrinsic motivation of pre-service primary school teachers for learning chemistry in relation to their academic achievement. *International Journal of Science Education*, *30*(1), 87–107.
Kelly, J. (2000). Rethinking the elementary science methods course: A case for content, pedagogy, and informal science education. *International Journal of Science Education*, *22*(7), 755–777.
Kirsh, I. (1995). Self-efficacy and outcome expectancy: A concluding commentary. In J. E. Maddux (Ed.), *Self-efficacy, adaptation and adjustment: Theory, research, and application* (pp. 341–345). New York, NY: Plenum Press.
Lardy, C. H. (2011). *Personal science teaching efficacy and the beliefs and practices of elementary teachers related to science instruction* (*Doctoral Dissertation*). University of California at San Diego.
Lawrance, G. A., & Palmer, D. A. (2003). *Clever teachers, clever sciences: Preparing teachers for the challenge of teaching science, mathematics and technology in 21st century Australia*. Canberra: Australian Government, Department of Education, Science and Training: Research Analysis and Evaluation Group.
Leonard, J., Barnes-Johnson, J., Dantley, S. J., & Kimber, C. (2011). Teaching science inquiry in urban contexts: The role of elementary preservice teachers' beliefs. *The Urban Review*, *43*(1), 124–150.
Levy, F., & Murnane, R. J. (2006). Why the changing American economy calls for twenty-first century learning: Answers to educators' questions. *New Directions for Student Leadership*, *2006*(110), 53–62.
Lewthwaite, B., Murray, J., & Hechter, R. (2012). Revising teacher candidates' views of science and self: Can accounts from the history of science help? *International Journal of Environmental & Science Education*, *7*(3), 379–407.
Logerwell, M. G. (2009). The effects of a summer science camp teaching experience on preservice elementary teachers' science teaching efficacy, science content knowledge, and understanding of the nature of science. Fairfax, VA: George Mason University.

Lucero, M., Valcke, M., & Schellens, T. (2013). Teachers' beliefs and self-reported use of inquiry in science education in public primary schools. *International Journal of Science Education*, *35*(8), 1407–1423.

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*(2), 153–166.

Magnusson, S., Krajcik, J., & Borko, H. (2002). Natures, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science teaching* (pp. 95–132). Dordrecht: Kluwer Academic Publishers.

Martin, M. O., Mullis, I. V., Beaton, A. E., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1997). *Science achievement in the primary school years: IEA's third international mathematics and science study (TIMSS)*. Boston: TIMSS International Study Centre, Boston College.

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science*. Retrieved from Boston College, TIMSS & PIRLS International Student Center website: http://timssandpirls.bc.edu/timss2015/international-results/

McDonald, C. V. (2010). The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science. *Journal of Research in Science Teaching*, *47*(9), 1137–1164.

McDonnough, J. T., & Matkins, J. J. (2010). The role of field experiences in elementary preservice teachers' self-efficacy and ability to connect research to practice. *School Science and Mathematics*, *110*(1), 13–23.

McKinnon, D. H., Danaia, L., & Deehan, J. (2017). The design of preservice primary teacher education science subjects: The emergence of an interactive educational design model. *Journal of Astronomy & Earth Sciences Education (JAESE)*, *4*(1), 1–24.

McKinnon, M., & Lamberts, R. (2014). Influencing science teaching self-efficacy beliefs of primary school teachers: A longitudinal case study. *International Journal of Science Education, Part B*, *4*(2), 172–194.

Menon, D., & Sadler, T. D. (2016). Preservice elementary teachers' science self-efficacy beliefs and science content knowledge. *Journal of Science Teacher Education*, *27*(6), 649–673.

Mulholland, J., Dorman, J. P., & 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.

Mullis, I. V., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. Boston: TIMSS & PIRLS International Study Center. Boston College.

Nillson, P., & Loughran, J. (2011). Exploring the developing of pre-service science elementary teachers' pedagogical content knowledge. *Journal of Science Teacher Education*, *23*(7), 699–721.

NSW Government – Department of Education. (2015). *2015 teaching workforce supply and demand*. Sydney: NSW Department of Education.

NSW Government – Department of Education and Communities. (2011). *2011 workforce plan – for school teachers in NSW public schools*. Sydney: NSW Government – Department of Education and Communities.

O'Toole, J., & Beckett, D. (2013). *Educational research: Creative thinking and doing* (2nd ed.). Melbourne: Oxford University Press.

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

Palmer, D. (2007). Practices and innovations in Australian science teacher education programs. *Research in Science Education*, *38*(2), 167–188.

Palmer, D. (2011). Sources of efficacy information in an inservice program for elementary teachers. *Science Education*, *95*(4), 577–600.

Palmer, D. H. (2006). Sources of self-efficacy in a science methods course for primary teacher education students. *Research in Science Education*, *36*(4), 337–353.

Richardson, G. M., & Liang, L. L. (2008). The use of inquiry in the development of preservice teacher efficacy in mathematics and science. *Journal of Elementary Science Education*, *20*(1), 1–16.

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

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*, 1–28.

Sang, G., Valcke, M., van Braak, J., Zhu, C., Tondeur, J., & Yu, K. (2012). Challenging science teachers' beliefs and practices through a video-case-based intervention in China's primary schools. *Asia-Pacific Journal of Teacher Education*, *40*(4), 363–378.

Sarikaya, H., Çakiroglu, J., & Tekkaya, C. (2005). Self-efficacy, attitude and science knowledge. *Academic Exchange Quarterly*, *9*(4), 38–42.

Schibeci, R. A. (1984). Attitudes to science: An update. *Studies in Science Education*, *11*, 26–59.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalised causal inference*. New York, NY: Wadsworth Cengage Learning.

Skamp, K., & Mueller, A. (2001). A longitudinal study of the influences of primary and secondary school, university and practicum on student teachers' images of effective primary science practice. *International Journal of Science Education*, *23*, 227–245.

Slater, S. J., Slater, T. F., & 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.

Smith, G. (2014). The impact of a professional development programme on primary teachers' classroom practice and pupils' attitudes to science. *Research in Science Education*, *45*(2), 1–25.

Sunger, M. (2007). *An analysis of efficacy beliefs, epistemological beliefs and attitudes towards science in preservice elementary science teachers and secondary science teachers (Doctoral dissertation)*. Çankaya/Ankara, Turkey: Middle Easter Technical University.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55.

Thomson, S., Hillman, K., Wernet, N., Schmidt, M., Buckley, S., & Munene, A. (2012). *Monitoring Australian year 4 student achievement internationally: TIMSS and PIRLS 2011*. Melbourne: Australian Council for Educational Research.

Thomson, S., Wernet, N., Underwood, C., & Nicholas, M. (2008). *TIMSS 2007: Taking a closer look at mathematics and science in Australia*. Melbourne: Australian Council for Educational Research.

Tschannen-Moran, M., & Hoy, A. W. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*(7), 783–805.

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

Tytler, R., Osborne, J., Williams, G., Tytler, K., & Clark, J. C. (2008). *Opening up pathways: Engagement in STEM across the primary-secondary school transition*. Canberra: Australian Department of Education, Employment and Workplace Relations.

Tytler, R., Smith, R., Grover, P., & Brown, S. (1999). A comparison of professional development models for teachers of primary mathematics and science. *Asia-Pacific Journal of Teacher Education*, *27*(3), 193–214.

Van Driel, J. H., & Berry, A. K. (2010). The teacher education knowledge base: Pedagogical content knowledge. In B. McGraw, P. L. Peterson, & E. Baker (Eds.), *Third international encyclopaedia of education* (pp. 656–661). Amsterdam: Elsevier.

Velthuis, C. (2014). *Collaborative curriculum design to increase science teaching self-efficacy* (*Doctoral dissertation*). University of Twente, Enschede.

Velthuis, C., Fisser, P., & Pieters, J. (2014). Teacher training and pre-service primary teachers' self-efficacy for science teaching. *Journal of Science Teacher Education*, *25*(4), 445–464.

Watters, J., & Ginns, I. (2000). Developing motivation to teach elementary science: Effect of collaborative and authentic practices in preservice education. *Journal of Science Teacher Education*, *11* (4), 301–321.

Weiss, I. (1994). *A profile of science and mathematics education in the United States*. Chapel Hill, NC: Horizon Research, Inc.
Williams, D. M. (2010). Outcome expectancy and self-efficacy: Theoretical implications of an unresolved contradiction. *Personality and Social Psychology Review*, *14*, 417–425.
Yilmaz-Tuzun, O., & Topcu, M. S. (2008). Relationships among preservice science teachers' epistemological beliefs, epistemological world views, and self efficacy beliefs. *International Journal of Science Education*, *30*(1), 65–85.
Yin, R. K. (2003). *Case study research – design and methods* (3rd ed.). London: Sage Publications.