

The Design Of Preservice Primary Teacher Education Science Subjects: The Emergence Of An Interactive Educational Design Model

David H. McKinnon, Edith Cowan University, Australia

Lena Danaia, Charles Sturt University, Australia

James Deehan, Charles Sturt University, Australia

ABSTRACT

Over the past 20 years there have been numerous calls in Australia and beyond for extensive educational reforms to preservice teacher education in the sciences. Recommendations for science teacher education programs to integrate curriculum, instruction and assessment are at the forefront of such reforms. In this paper, we describe our scholarly action–research approach to the teaching of science and science–method subjects to Australian preservice primary-school teachers in the state of New South Wales. We present an interactive educational design model founded on a solid theoretical literature base that incorporates Pedagogical Content Knowledge as an integrative mediating framework and which drives students’ interactions with the elements of the design model. The results from our mixed-methods study suggest that the approaches adopted through two extended vignettes show significant increases in preservice teachers’ competence and confidence. Together, the qualitative and extensive quantitative data suggest participants obtain a newly developed sense of enthusiasm for science and an understanding of the role that it can play in the primary- school curriculum. The data provide strong evidence that the approaches being called for in some of the earlier reforms and most recently by Bybee (2014) are effective

Keywords: Astronomy Education Research; Teacher Education; Pedagogical Content Knowledge; Science Teaching Efficacy

 Successfully engaging elementary-school students in science requires competent and confident teachers capable of implementing engaging pedagogies (Appleton, 2003; Bybee, 2014; Tytler, Osborne, Williams, Tytler & Cripps–Clark, 2008), but many perceive themselves to be neither (Appleton, 2003; Bellocchi, Ritchie, Tobin, King, Sandhu & Henderson, 2014). Effective preservice elementary- teacher education, therefore, is one important way to build both of these attributes in our future elementary teachers. Attempts to address the issue of *competence* in many preservice teacher education subjects in both Australia and overseas involve the science content being taught by Science Faculty in isolation from the pedagogies or instructional methods that they will require to teach it. That is to say, the science is taught by those who lack the pedagogical insights of what to do in the elementary-school classroom. Education Faculty, in the science methods subjects, typically focus on how to teach specific and different science topics each week in an attempt to cover the broad spectrum of the curriculum’s demands and to provide experiences for the pre–service teacher to employ a range of pedagogical approaches that can be used to teach the science content. Bybee (2014) criticizes this separated approach not the least because the audience do not have *any* specialized training in *science* nor indeed have they had much in their brief exposure to the broad science content of the elementary-school curriculum. They thus lack much of the science content knowledge that they will be expected to teach (Appleton, 2007; Goodrum & Rennie, 2007; Tytler, 2007). The consequence is that Australian elementary-school teachers do not feel themselves to be *competent* and thus they lack *confidence* when confronted by the demands of the science curriculum that they have to teach. One result is that science education in the elementary school is often avoided or ignored, even although it is mandated (Angus, 2004). Indeed, the 2015 TIMMS data reports that the *average* amount of time spent doing science in the elementary school is 57 minutes (Thomson, Wernert, O’Grady &

Rodrigues, 2016). Overall, the performance of Australia's Year 4 students is flat-lining while the performance of students in other countries is improving (Thomson et al., 2016). Teachers are key to improving this situation.

Bybee (2014) highlights the need for an *integrated* approach to curriculum, instruction and assessment in preservice science education subjects. He notes that such an approach would be very different to the disjointed or fragmented approach of how science is currently and typically covered by teacher education institutions. The more traditional approaches employed by many teacher education institutions around the world tend to employ a fragmented approach to the science content, one which is much criticized by Bybee (2014). That is to say, a typical approach focuses on pedagogical techniques on how to teach particular concepts with little developmental continuity in the science content. For example, this week is 'little beasties', next week is 'the weather' followed by 'push and pull' and perhaps the week after 'the Earth and Moon' (Bybee 2014) and so on.

In contrast, Bybee (2014) claims that the different and integrated approach would involve deep, intense experiences that combine disciplinary knowledge with scientific practices through students undertaking investigations. In the Australian context, Lawrance and Palmer (2003) and Palmer (2007) identified a number of innovations in elementary-school teacher education at both the program and the subject level in a sample of tertiary institutions some of which foreshadowed Bybee's (2014) call for the integrated approach. Innovations included: developing positive attitudes by emphasising the use of motivating "hands-on" activities and the use of real-life examples; the integration of science with mathematics; the use of interdisciplinary approaches that integrate science and mathematics with other curriculum areas; the use of problem-based learning to give students some control; the integration of content and pedagogy through the use of themes; the use of constructivist approaches to investigate real-life problems to build positive attitudes; and, linking theory with practice through incorporating visits to schools to practise the science pedagogies about which they have been learning. Lawrance and Palmer (2003) noted, however, that the institutions included in their study typically displayed only one of these innovations, which was often instigated and maintained by only one or two motivated staff members. Consequently, educational-design development and succession planning were problematic.

Many key researchers and writers in the field also highlight the importance of Pedagogical Content Knowledge (PCK) in preservice teacher education subjects (e.g., Appleton, 2003, 2007; Hofstein, Eilks & Bybee, 2010; Lawrance & Palmer, 2003; Tytler, 2007). This construct appears to be missing from the innovations listed above. It could be argued that PCK is the key construct that melds many of these innovations within a constructivist approach to the learning and teaching of science for preservice elementary-school teachers. This melding may provide important synergies to render science-methods subjects accessible, engaging and understandable to produce elementary-school teachers who are more motivated to teach science. Research indicates that PCK can be developed through teacher education interventions (Franke, Carpenter & Levi, 2001; Saxe, Gearhardt & Nasir, 2001).

Moreover, we would argue that there is a pressing need for the inclusion of detailed and replicable descriptions of science interventions in the science education research literature to serve the dual purposes of facilitating replication studies in alternate contexts and the goal of effecting lasting, positive changes to tertiary science education practices globally. Deehan (2016) articulated the need for more detailed descriptions of science interventions when he highlighted the unknown relationship between the use of innovative practices reported in the literature (Lawrance & Palmer, 2003) and student outcomes in a review of the science teaching efficacy belief literature. It appears that we know what works but not how successful, innovative pedagogical approaches can be implemented in different educational contexts. Indeed, concepts such as cooperative learning, constructivism and inquiry learning are often presented in the theoretical domain with limited discussion of practical pedagogies (e.g., Abrams, Southerland & Evans, 2008; Settlage, 2007). It could be argued further that results-driven-publication processes place unintentional limits on the intervention descriptions offered by researchers (Cantrell, 2003; Utley, Bryant & Moseley, 2005). However, in recent years researchers have attempted to articulate replicable science interventions in academic publications (Lewthwaite et al., 2012; Northcutt & Schwartz, 2013). If successful innovative interventions are to become mainstream inclusions for the preparation of pre-service teachers in science education, deep descriptions of the interventions and the accompanying educational design processes as well as their effects need to be encouraged and presented within the research literature.

The purpose of this paper is thus fourfold. First, an interactive educational design model (IEDM) is presented that emerged from our scholarly interrogation of the message systems of education, curriculum, pedagogy and evaluation (Bernstein, 1971). Second, the IEDM-in-action is illustrated using two extended vignettes of the science subjects to illustrate the interactions amongst the design elements and to demonstrate the mediating role played by PCK. Qualitative data collected from students as they undertook their journey in these science subjects is integrated within the vignettes to illustrate the effect of the design. The model contains many of the innovations outlined by Lawrance and Palmer (2003). Importantly, it incorporates PCK as an integrative mediating framework that drives all students' interactions with the design elements. After 2008, the IEDM was employed in two science subjects in a Bachelor of Education degree at an Australian university for elementary-school preservice teachers all of whom were participants in the participatory action research (McIntyre, 2007) conducted and for which ethical clearance was granted (Protocol Number 2006/122).

Third, quantitative data collected over a ten-year period are presented to illustrate the impact of the implementation of the IEDM in terms of the Appleton (2003) criteria of *competence* and *confidence*. Two sources of quantitative data are presented: a shorter version of the Astronomy Diagnostic Test (CAER, 2002, 2004) comprising 15 items that are mapped to the content requirements of one component of the elementary-school Australian Science Curriculum and which is administered on both pre- and post-subject occasions separated by an interval of 14 weeks; and, a slightly modified form of the Science Teaching Efficacy Belief Instrument-B (Enoch & Riggs, 1990).

Fourth, the discussion examines the implications for the future design of tertiary elementary-school teacher education preparation programs in science education in light of the changes called for by Bybee (2014) and Tytler (2007) within the new Australian Science Curriculum. The presentation of an IEDM, and two extended vignettes, will serve as a replicable model of student-centered tertiary science education. It thus helps to bridge the divide between preservice elementary-school teacher education research and mainstream science teaching practices in tertiary institutions.

BACKGROUND AND CONTEXT

In 2004, we were charged with delivering two science subjects at very short notice. With no time to make changes, we delivered what had been done over a number of years. In this mode of delivery, it rapidly became clear to us that our students were afraid of science, were not engaged by it, and chose to talk about their social lives rather than engage with the science topics they encountered in class. Moreover, less than 10% of them read the assigned subject materials and their attempts at the assessment items left much to be desired. A major evaluation at the end of the semester of what had been achieved (i.e., little), led us to formulate a participatory action-research project that we implemented in 2005. In moving from the old subject design to the new one, we asked two key questions: "*What characteristics should our students possess on finishing each of these subjects?*" and "*How would we know?*" In defining these characteristics, we arrived at a new assessment frame that we then backward mapped to define the learning experiences students should encounter.

We were also motivated to generate subjects that drew on a number of approaches that research had demonstrated led to engagement and high performance. These included cooperative learning (e.g., Johnson, 1994), problem-based learning (e.g., Duch, Groh & Allen, 2001) and, mastery learning (e.g., Bloom, 1971) all of which were designed to interact in ways that developed in our students a sense of *competence* with science and who were sufficiently *confident* to teach it. We also decided that "less was more" in the sense that we restricted the science content to two domains, astronomy and environmental science, rather than cover a new topic each week, which had been the previous modality. In the second semester, 2005, we assumed control of the second science methods subject that was designed to extend both students' pedagogical repertoire and their knowledge of scientific content.

Our university had adopted the Bologna Model for Higher Education (Van Damme, 2001) to inform the workload that was required of students. Delivery of our two subjects was face-to-face in a one hour mass-lecture format followed by a two-hour practical class in a room that emulated what an elementary-school classroom is like. That is to say, it had moveable tables, a wet area with three sinks and two small storage rooms. We refer to these two-hour classes as a "practical class", which had a maximum of 24 students divided into six cooperative learning groups. We built a workload that would take a successful student eight to nine hours each week to complete at home or in the library.

This requirement, in conjunction with the face-to-face classes met the requirements of the Bologna Model of 144-166 hours of study over the 14-week semester.

A small internal teaching grant provided six networked computers in the practical classroom: one for each table group. Students were also encouraged to bring their laptop computers to the practical classes, which also had access to the university network through a WiFi access station installed in the room. Such infrastructure allowed students to access the Internet and subject materials as required. Our intention was that the computers would be used as “learning tools” by the students.

Interrogation of the impact of these two redesigned subjects with the two different cohorts of students in 2005 led us to modify both for the 2006 offering. One major modification was to the mastery learning and assessment frames. This modification involved students being awarded the mark that they achieved on the *first* occasion of their submission of an assignment. If they had not demonstrated mastery over all components, they could choose to re-submit, and could continue to do so, until mastery had been achieved in which case they became eligible for a *passing* grade in the subject. This overcame the tactic of some high achieving students continually re-submitting their attempts to get higher and higher marks even although they had achieved mastery on the first submission.

In undertaking the learning experiences that had been mapped to the assignments, it was clear that many experienced difficulty in understanding the science concepts and of understanding why the subject format had been changed so dramatically. Consequently, in 2006 we implemented an approach that sought feedback after *every* class. This took the form of a one-minute paper that asked five questions: *What worked for you? What could be improved? What did you learn? What do think you now need to learn?* and, *Can you give five words to describe your feelings about what happened this week?*

Initially, this formative feedback was collected on a hand-written paper template. Analysis of these was tedious and so we progressed to using an online survey form that enabled students to complete the questions before leaving the practical class. This allowed us to analyse rapidly their formative feedback and collaboratively to discuss what some of it actually meant. At the start of the lecture during the following week, we presented our analyses and conclusions together with what we were going to do to address particular problems/situations that week. In some cases, we set small cooperative learning tasks to address issues that students could tease out amongst themselves (e.g., why we were now doing science *this way*).

After the very powerful gains in students’ content knowledge achieved in the first year of the project, we were somewhat surprised and disappointed with the much lower gains in 2006. Consequently, in 2007 we required that students directly deal with *all* of their alternative scientific conceptions during their “micro-teaching” experiences planned for each week throughout the first subject. These are more fully described in the first vignette below. In addition, further interrogation of the 2006 outcomes led us to introduce, in 2007, the Science Teaching Efficacy Belief Instrument–B and to use the Personal Science Teaching Efficacy (PSTE) scale as a *de facto* measure of students’ *confidence* to teach science. In a different sense, the Science Teaching Outcome Efficacy (STOE) scale became a *de facto* measure of how they perceived the subject design changes that we had introduced impacted on themselves.

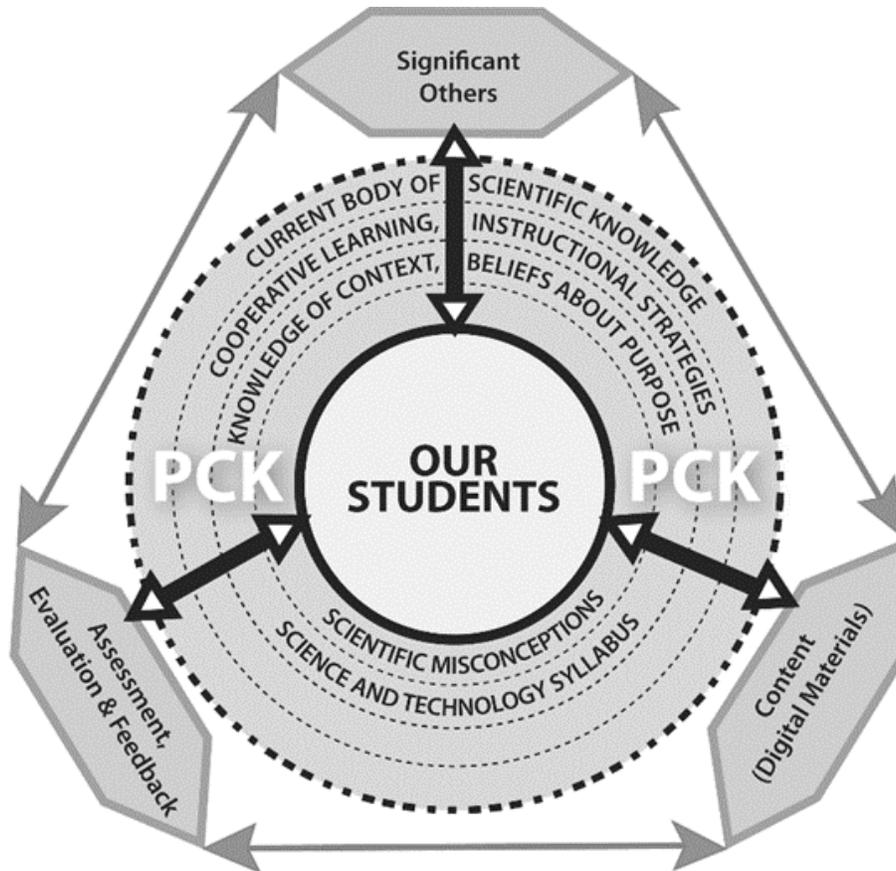
In early 2008, we struggled to understand what seemed to be motivating our students to become curious, to appreciate science, to see its power as an integrative frame for other curriculum areas, and leaving them wanting to design and implement engaging learning experiences for their own pupils when they arrived in schools as teachers. In this scholarly approach, described by Shavelson, Young, Ayala, Brandon, Furtak and Ruiz–Primo (2008), we pursued an almost constant interrogation of our practice. As we investigated the apparent interactions amongst the various components of the designs of the two subjects, an explanatory model emerged.

The model that we describe below represents a clearly defined method of developing preservice teachers’ science PCK in ways that are applicable to science teaching practice in the elementary school. Continued interrogation of the model in subsequent years has revealed its stability and power. Our analyses have allowed us to anticipate the conceptual difficulties that students are likely to encounter as well as their evolving concerns as the semester unfolds and to address these to help allay their fears.

The Interactive Educational Design Model

The Interactive Educational Design Model (IEDM) that emerged in 2008 is presented below in Figure 1. It combines a number of well-researched educational-design frameworks in pedagogy and assessment. Elucidation of the model below illustrates the ways in which the authors now intentionally craft a profound personal learning journey for the elementary-school teacher education students that requires them to extract, from their in-class educational experiences, the knowledge and skills required to create inspiring science classrooms.

Figure 1. The Interactive Educational Design Model (PCK= Pedagogical Content Knowledge)



The model comprises three key design elements located at the vertices of the model: *Content*; the *Assessment/Evaluation/Feedback* system; and, *Significant Others*. These elements are described first before considering their interactions. The IEDM embeds students within the construct of Pedagogical Content Knowledge (PCK) and its development as the mediating domain that deeply influences all interactions between students and the design elements. The location of students at the centre of the IEDM signifies the student-centered nature of the approach.

The Content

This design element of the IEDM facilitates students' access to all of the content and the resources developed for the subjects. The materials are provided at the commencement of the semester in a digital format both via USB and through the Learning Management System used by the university. The materials are organised into multiple folders

with access facilitated by extensive hyperlinking from within the collection of PowerPoint lectures and Word-based tutorial documents as well as to other applications such as spreadsheets, movies and software.

Delivering the content in this way is designed to achieve two major aims. First, students have access to all of the materials from the outset of the semester of study. Second, it allows the teaching teams across multiple campuses to achieve a high degree of implementation integrity. The latter issue is of importance given that multiple instructors are employed across our multi-campus university to deliver the subjects and to evaluate the assignment products in a consistent way. Students' reactions to this level of organisation and distribution is extracted from data collected in the Evaluation and Feedback System. Many examples are presented within the two vignettes.

The Assessment/Evaluation/Feedback System

These three components constitute one design element in the model and are presented separately in this section. Nonetheless, they are constructed to interact in powerful and supportive ways. The vignettes that follow the description of the IEDM illustrate the interactions amongst them.

Assessment. The assessment domain is constructed by asking ourselves the question “What would be the characteristics of the student who has *successfully* completed this subject?” This then begs a further question of “How would we *know* that a student possesses these characteristics or has acquired them during the subject?” The assessment domain is thus defined by the process of Backward Mapping (Shavelson et al., 2008) with the early assignments each addressing a particular component of PCK. In defining these student characteristics, the assignments are developed in ways that allow instructors to be confident the students have acquired the knowledge, the skills and the attitudes and values, which the instructors wish them to acquire.

To help the students achieve the outcomes, assessment rubrics contain clearly defined performance criteria and are supplied to the students within the digital materials. The standards within the rubrics are specified in such a way that the student is in absolutely no doubt of why they were assessed at a particular performance level. Words like *Satisfactory*, *Good*, *Very Good* and *Excellent* in relation to any component of a report or essay are an anathema to students who cannot decode these words in any meaningful way. Nor are they useful to instructors who are left with only their own qualitative experience and judgement for making any call on the differences between *Good* and *Very Good* or *Excellent*.

Consequently, the authors employ a criterion-referenced assessment framework with a set of clearly specified performance outcomes for each of the components of each assignment. Specific criteria, or standards, for each component also present clear structural guides for student performance from *Pass* through *High Distinction*. These criteria have the added benefit of increasing the reliability of assessment marking for the instructors. The joint effect of these analytic rubrics is that students understand what is expected of them and the standards at which *they* can choose to deliver their attempts at the assignment in order to be deemed eligible to pass the subject. The construct of PCK progressively and cumulatively mediates students' interactions with the sequence of assignment items, their engagement with the curriculum content, focuses their attention on each of the sub-components of PCK in a developmental way and, assists them to understand the role that PCK will play in their design and delivery of science learning experiences for their future pupils.

Feedback. Feedback occurs at a number of levels on multiple occasions during a semester: from students to the instructors, from instructors to the student, and from student to fellow student. Students provide formative feedback to the instructors each week. They, in turn, provide both formative and summative feedback to the students on a weekly basis and on the occasions when assignments have been tendered for assessment. Students provide feedback to each other as they develop their cooperative learning products and when micro-teaching events happen (Cochran et al., 1991; Niess, 2005). Summative feedback in relation to the delivery of the subject is provided by the students to the instructors at the completion of the semester of study both through the formal University online system and through an extended paper-based questionnaire. At the completion of the semester, the summative feedback provided by students coupled with the formative feedback collected each week are interrogated by the instructors in a scholarly

way. Interventions can thus be implemented in at least two ways: “on-the-fly” each week to deal with any barriers to the students’ learning, or in a structural way to provide for more major changes to a subject during its next offering.

The use of “one–minute feedback papers” obtained at the end of each class provides solid evidence of the way in which the instructors are concerned that no student is left behind. This formative feedback mechanism was introduced to elicit information from students when it became clear that, despite our early efforts, some students could not disentangle the “message” from their “feelings” about what was happening. Consideration of this form of feedback thus allows the instructors to undertake deeper reflections about their teaching and to develop interventions that are implemented in the following lecture where students are presented with an analysis of the feedback received and the actions that will be undertaken that week. Analysis also allows the instructors to identify learning barriers and conceptual problems that can be dealt with expeditiously before they become a problem.

The support mechanisms for students’ development as cooperative learners led to the creation and implementation of a *Cooperative Learning Evaluation Feedback* instrument to identify emerging difficulties in any interactions within their groups. Once problems are identified, the instructors act as “counsellors” to address the issues within a group but take care not to single out individuals. Members of a problematic group are directed to relevant literature supplied to them in the digital materials on how to function more effectively, for example, a short paper on *Assertive Communication* and how to implement it (Hoffman, 1978). Discussion then follows that is led by the instructor on how they might deal with *their* problem. The group’s progress is monitored together with effective forms of communication being modelled and positive feedback provided as skills are developed and demonstrated. This support is important because, as teachers, they will be required to collaborate with their peers as well as to provide assistance for their future pupils to do the same. When our students come to conduct experiments and teach the content to each other, a *Structural Reflection and Feedback* qualitative instrument helps focus their attention on what worked and what did not. This instrument asks students to provide qualitative feedback to their peers guided by our questions in the instrument to focus their attention on key aspects of good science lessons, e.g., the introduction, linking to previous work, identification of alternative conceptions, trying to disprove each other’s conceptions through experiment, assessing the explanations given by the “teacher”, linking forward to what is to follow, seeking feedback on understanding etc. Thus, the instrument cannot be assessed in quantitative terms. In addition, it provides a formal assessable mechanism for them to provide feedback to the peer who had attempted to teach them, *viz.*, they have to provide the feedback in order to get credit for it.

Subject Evaluation. Instructors at our university are required to evaluate each subject at the end of each teaching semester. This process is achieved through asking students to respond to a number of core (11) and optional (29) items provided by a central body within the university charged with evaluating the delivery of all subjects. The core items are common across all subjects within the university. Instructors can choose additional optional statements from an item bank because of their relevance to aspects of a particular subject. All items require a response on a seven–point Likert Scale from *Very Strongly Disagree* through *Very Strongly Agree*. In addition, students may add written comments if they feel so moved. The average response rate for these anonymous surveys at our university is typically around 30%. However, in these two subjects it is always greater than 90%. Typical extracts from some of these summative subject evaluations are provided in the two extended vignettes below.

Significant Others

The students rapidly realise that their peers within their group, within their practical class and, more broadly, across their cohort constitute one aspect of the design element of *Significant Others*. The realization is rapid: indeed it happens within their first week of their first semester of study. They understand that they are dependent on the cooperative learning products that others within their group, their practical class and their cohort generate and all of which contribute to the production of the assessment items in part, or as a whole. For example, in writing an academic essay on alternative conceptions, they understand that everyone in the year group contributes to the literature base to which they are exposed in a jigsaw fashion (Slavin, 1996) where the extensive reading list has been divided amongst the entire cohort. Thus, they understand that they will cooperate with others to generate their assignment responses for assessment.

The instructors also play a role as significant others. They model for the students within their own particular practical class the behaviors that are required of them when they become teachers. They model how to: interrogate the scientific content; evaluate and comment on the appropriateness of instructional strategies to teach particular content; assess the learning outcomes to be achieved; mediate when difficulties arise in the social dimension; and, set realistic goals.

Finally, *Significant Others* also include those whom the students will, perhaps, never meet personally. These are the science education researchers to whose work they have been exposed, and the scientists whose work they will encounter and include in their classes when they become teachers.

The five key elements of cooperative learning drive students' understanding of the extent to which they are dependent on the work of significant others in their journeys to become teachers. The five elements are: face-to-face promotive interaction; positive interdependence; individual accountability; interpersonal and collaborative skills; and, reflection/group processing of all interactions (Johnson, Johnson & Smith 1998).

The Interdependence Amongst the Design Elements

The longer and lighter double-headed arrows in Figure 1 signify both the interdependency amongst the design elements and the dynamic and mutually supportive nature of their interaction. For example, when preparing an assignment, students are required to assimilate certain *Content* delivered in the digital materials and to interact with *Significant Others* in the form of their peers and instructors. Interaction amongst the design elements thus creates the environment for students to learn both the scientific and educational content, to appropriate a variety of instructional strategies, and to produce quality assessment responses throughout the subject.

The Mediating Role of PCK

We have operationalised the construct of PCK to comprise the layered and embedded components of alternative conceptions, context and beliefs about purpose, the Science and Technology Syllabus, cooperative learning and instructional strategies, and the current body of scientific knowledge (Cochran, DeRuiter, & King 1991; Grossman, 1990; Shulman, 1986). These mediate all interactions that the students are required to have with the design elements. These components are instrumental in developing the more complex construct of PCK as these preservice teachers begin to make decisions *within* each of these components and to move *amongst* them in response to the complex and unfolding learning context.

Each of these components of PCK is bounded by broken lines to indicate that they are not fixed. That is to say, knowledge of these components of PCK grows with time, with critical reflection, and with their expanding personal experiences in science education. They come to understand that they do not stand in isolation from one another. These components of PCK provide them not only with a framework within which to engage with the contents of the two subjects but also with a context to access a wider range of ideas and the significant others with whom to explore and discuss effective science learning and teaching.

Thus, these components of PCK mediate all students' interactions in varying ways with each of the three design elements. The short black double-headed arrows thus serve to indicate that the students interact with each of the three design elements in ways that are mediated by their interactions with the components of PCK. Most importantly, locating the students at the centre of the IEDM serves to indicate that the locus of control rests with them, that is, when they are studying the subject materials, communicating with their peers, instructors or others, or preparing assignment responses.

The Two Extended Vignettes

The vignettes presented here illustrate how the design elements of the model come together to form an integrated approach to the education of preservice elementary-school teachers whose role will include the teaching of science. They further illustrate how cooperative learning and backward mapping of the assessment tasks are instantiated within the learning experiences. The context for their journey in Subject 1 is a problem-based learning environment that is

constructed to address their limited scientific knowledge, and the many alternative conceptions they hold. Subject 2 progresses the educational outcomes developed in the first subject. These are enacted within an authentic project-based learning environment to extend students' mastery and professional orientation in readiness for their transition to becoming reflective practitioners and effective teachers of science in elementary schools.

The First Subject

Subject 1 is one of four that students experience in their first semester of a four-year Bachelor of Education (Elementary) degree leading to employment as a elementary-school teacher. This subject is largely a science content one but taught in a markedly different way to those criticized by Bybee (2014). We *integrate* the learning of the content with the pedagogies students will employ as teachers of science using elements of Pedagogical Content Knowledge as a developmental frame within a problem-based learning (PBL) environment.

The authors create the PBL environment at the outset of Subject 1 by administering a modified version of the Astronomy Diagnostic Test (ADT) (CAER 1999) to students at the first whole-group lecture of the subject. The 15 items are mapped to the outcomes of the elementary-school science curriculum and have been modified by the authors for use in the southern hemisphere. That is to say, pupils are expected to know the material contained within this version of the ADT on exit from elementary school. In addition, our students are invited to present the reasoning for their answers. This allows us to analyse the structure of their alternative conceptions. The ADT thus serves several purposes. It creates the preconditions for the PBL environment where *their* learning is the issue to be addressed. Their results shift their perception that it '... is only elementary-school science. I have been to high school and so I will know more than them' (GN). More importantly, in illustrating what they do not know, it defines what they will need to learn in the classes that follow (Hickey, Taasobshirazi & Cross 2012).

In considering the issue of *confidence to teach science*, the work of Bandura (1977) has much to offer where the construct of "self-efficacy" is defined as a general anticipation about future events (the outcome domain) based on previous personal life experiences (the personal domain). Enochs and Riggs (1990) developed an instrument that attempts to measure science teaching efficacy in these two domains. One, Personal Science Teaching Efficacy (PSTE) is "a teacher's confidence in his or her own teaching abilities" and the second is the Science Teaching Outcome Expectancy (STOE) as "a teacher's belief that student learning can be influenced by effective teaching" (Ramey-Gassert & Enochs 1990). The instrument was later modified slightly for pre-service teachers called the Science Teaching Efficacy Belief Instrument-B (Riggs & Enochs, 1990). The students are asked to complete the Science Teaching Efficacy Belief Instrument-B (STEBI-B). On pages 35-36 reference is made to how it is used within both subjects.

In self-marking their attempt at the ADT in the following practical class, our students are able to identify their knowledge gaps in the astronomy component of the syllabus and gives them almost immediate feedback concerning the many alternative conceptions they hold (Hickey et al., 2012), for example, that it is the Earth's shadow that causes the phases of the Moon, or the Earth's distance from the Sun that causes the seasons. Their marks typically range from zero to a maximum of five out of 15 possible marks with a mean of around 1.5 each year.

We then confront them with the Science Syllabus outcomes for Grades 3–6 against which the 15 items are mapped. The impact on the emotional climate of the class is both immediate and profoundly negative (Bellocchi et al., 2014). They are asked to record their personal reactions in a Word document on their personal, or tablet, computers. They use words such as 'horrified, shocked and stupid' in reporting their feelings. We then ask them to form groups of four and to share both these reactions and their score with their new table-group peers. We also ask them to record the varying reactions expressed by their peers. By this stage they have almost completed the homework for the first low-stakes assignment that requires them to write a short reflective essay about what they know, what they will need to know if they are to become elementary-school teachers, and to map their lack of knowledge against the Science Syllabus requirements in this content area. In this process, they begin to navigate the Science Syllabus document. They also discuss their feelings about both their own result and those of their group before presenting a "conclusion".

In considering these results, they clearly understand that there is a problem. They articulate it on both the first formative feedback occasion using the one-minute-feedback instrument and in their first assignment responses: '[T]o teach science, knowledge is needed... more knowledge than what I have' (Anon); 'If I don't know the content then how will I know how to teach it?'(LD); and, '[T]he test helped me to understand just how much I need to learn' (Anon).

We also analyse how they feel from the five stream-of-consciousness words used by them in the one-minute-feedback papers. Words such as 'angry, anxious, confronting, daunting, embarrassed, shocking, and stupid' are used with high frequency. The anger that they express relates to the fact that they 'don't remember being taught this stuff at school' (many). Over 80% of these words can be classified as negative and reveal the extent to which they are shocked and stressed yet, at the same time, also influenced and motivated by the revelation that they know little. This is illustrated by further words interspersing the negative ones above that constitute around 10% to 15% of the total and which may be classified as positive: 'eager to learn more, enjoyable, exciting, learning opportunity and motivating'. In a relatively brief period of time, the climate changes from a sense of personal failure, to an intermediate one of relief that each individual is not alone, to one involving anger yet balanced by a sense of personal motivation especially when the instructor emphasises that the problem is even more general than what they have just experienced (Bellocchi et al. 2014). The endemic nature of these alternative conceptions becomes apparent in their preparation for second assignment that begins immediately afterwards.

In this first class, we further develop the potential for increasing their motivation through exposing them to the extensive research literature on alternative conceptions possessed by pupils, teachers and preservice teachers. In this exercise, and consistent with collaborative learning principles, the task of reading and analysing the body of literature is too great for any one individual to execute in the time available (18 days) to produce the second assignment. Thus, we employ a variety of cooperative learning strategies including jigsaw, roundtable, think-pair-share, and numbered-heads-together (Slavin, 1996), to demonstrate that alternative conceptions are widespread. In the process, they come to understand that their lack of knowledge is not unusual.

For homework, each student is required to read one supplied journal article dealing with the alternative conceptions of a particular population (teachers, preservice teachers, elementary-school students) and to locate and read a second one on an associated topic. In order to scaffold their reading of the research papers, they respond to a set of six questions that focuses their attention to elicit relevant information pertinent to their second assignment. For homework, they summarize these responses to the two papers before meeting with a partner who has read and answered the same questions on two different papers covering the same population. This think-pair-share exercise requires the pair to come to the second class with a digital summary of their findings covering the four papers. Each group then employs a roundtable approach to understand each pair's findings for each of the two populations covered within their group. A numbered-heads-together strategy follows allowing the members from the various groups to share their findings on the alternative conceptions possessed by the three populations: school students, preservice teachers and in-service teachers. On return to their own group, a form of roundtable that we have called "cascading roundtable" is employed to share what each individual has learned from the other groups. This form of roundtable is designed to elicit for discussion, without repetition, all of the information supplied by the various groups. In this process, we demonstrate the value of the cooperative learning strategies being employed, their pedagogical value, and make this explicit for the students.

These processes provide the scaffolding necessary for them to understand a large body of literature that deals with the background reading necessary for the production of their second assignment: an academic essay on the non-scientific conceptions held by the three populations. Extracts from this second assignment demonstrate that this literature has a deep impact on their motivation. Many state that they will have to identify and address their own alternative conceptions first so that they will not teach these to their pupils, for example, 'I am shocked by the information that teachers teach alternative conceptions' (LD). Moreover, they are highly motivated to do something about "their problem" as exemplified by the following extract from one essay.

I feel that this diagnostic test was both revealing and confronting. It is scary to think that after 13 years of schooling I still have no idea about something as simple as the relationship between the Earth, Sun and

Moon. It is from this realization that the motivation comes to create an exciting and memorable science experience for students in schools (EC).

These highly personal experiences provide a framework for the groups to construct a learning program that is submitted as a cooperative group assignment. The third assignment requires them to sequence the investigations that will address their alternative conceptions, map each of these against the syllabus outcomes, and identify who will be responsible for teaching each one to their small group. The activities they *have* to choose are to meet their collective content–knowledge deficits. The assignment gives them a real purpose for the learning of the astronomy content that is to follow where they will experiment with instructional strategies to teach each other. This assignment drives their interactions with the mandatory science curriculum documents, and with the terminology therein, as they begin to grapple with sequencing their investigations in some sort of logical way with help from their instructors. The groups incorporate relevant projects into their learning program from the digital compendium of projects provided by the authors. From these projects, we have deliberately removed any hints on *how* to teach them so that they can experiment with various instructional strategies that engage both them and their peers and later, perhaps, after reflection, the pupils in their classes (e.g., explicit teaching, conceptual change, cooperative learning strategies, guided and open discovery, modelling, and interactive teaching).

In engaging with these investigations led by a peer, a beginning understanding emerges of the importance of this battery of instructional approaches to the effective teaching of science. This emerging grasp of the importance of the strategy was illustrated when one student said after experiencing the project on the cause of the seasons taught by one of his peers using explicit instruction ‘[B]ut I wouldn’t teach it this way’ (MR). When asked why, he responded in ways that indicated his developing understanding of the issues by pointing out the need to break the task into ‘manageable chunks’ over a series of lessons in order to ‘scaffold the learning of his future pupils’ (MR).

In the six weeks during which the group members attempt to teach the projects to their peers, they employ a Jigsaw II strategy (Slavin, 1996) and role cards. Each group member is required to lead the teaching of two of the projects. For any single investigation, one student assumes the role of teacher and thus has to become more expert on the topic in order to teach it to his/her peers in the small group, while the others act as learners. These roles are swapped for other investigations. This entire process requires everyone to be involved working in their different roles during the class with students as teachers, students as learners, and the instructors as facilitators and mentors. These sessions are almost chaotic as each group undertakes different investigations according to their group’s needs and for the sequence they have adopted. Nonetheless, the classes are characterised by continuously high levels of task engagement.

Formative feedback illustrates the extent to which they are engaged: ‘I loved the Jigsaw activities – everyone brought something different’ (NG); and, ‘We had a cooperative day where we got together and did it’ (RO). As they engage with the content, there are many, many “Ah Ha!” occasions when students come to understand such phenomena as the phases of the Moon or the seasons using conceptual change, modelling and guided discovery approaches.

The change in students’ normal pre-intervention alternative conception of “distance from the Sun” as being the cause of the seasons illustrates various processes that seem to happen almost simultaneously. The students’ alternative conception is relatively easily “explained” in our Australian context where being “close the fire/radiator/heater makes me feel warm/hot and moving further away lets me cool down”. This generalisation of the inverse square law to cover a different phenomenon has been present in all cohorts and held by over 95% of students. Their alternative conception is destabilized by asking the question “In December/ January here in Australia it is very hot. What is the weather like in Denver/Moscow/London at the same time?” They model the position of the Earth as it orbits the Sun over the course of a year and examine the lengths of the shadow cast by tin-tack at local midday. They measure the length of the shadows and are asked “Why is it so?” They understand from the modelling that the altitude of the Sun depends on two things: the location of the Earth in its orbit and the tilt of the Earth’s axis of rotation, which remains constant. They model the effect using a torch taped to a meter ruler so that it remains the same distance from a Lego person at the foot of the ruler. They see the varying shadow lengths cast by the Lego person and relate these to the altitude of the Sun above the horizon. They draw on graph paper the envelope of the bright beam at the greatest altitude and the lowest altitude. They see that the same amount of energy is spread over a greater area in winter compared with summer. They calculate the ratio of these two areas. They also use planetarium software to calculate the length of time that the

Sun is above the horizon on the summer solstice and on the winter solstice. They understand that the length of time that the “big yellow heater in the sky” is also partly responsible for the summer/winter temperature difference. Thus, the Earth-orbit model and shadow length of the tin-tack coupled with the torch model and planetarium software allows them to accommodate the scientific explanation for the difference in the seasons.

In the process of accommodation, they construct highly personal mental models to explain various phenomena. For example, one student explained the apparent movement of the Sun in the sky during the course of the year based on the peeling of an orange. The sharing and explanation of this idiosyncratic mental model excited both the originator and her group as the abstract idea for the cause of the seasons was made concrete, *viz.*, varying day length and varying angle of incidence of insolation due to the varying altitude of the Sun over the course of a year are all caused by the axial tilt of the Earth as its orbit around our Sun.

Students’ reactions to the learning and teaching experiences that are occurring are formalized in the Structural Reflection and Feedback instrument we developed to focus their attention on the interactions of the components of PCK (knowledge of: content; students’ prior conceptions; the syllabus; and, appropriate instructional strategies) they have been encountering so far within the subject. Feedback obtained during these sessions reveals students’ changed feelings (e.g., ‘active, beneficial, inspired, collaborative, engaging, enjoyable, informative, intense, interactive, interesting, powerful, productive, supportive, thought-provoking’). The balance of negative and positive words elicited in the one-minute feedback dramatically shifts to more than 90% positive and less than 10% negative by the end of Week 5, which is the second week of these teaching explorations. They now come to class with smiles on their faces rather than frowns. Our conversations with them reveal that they seem to “get it”.

Other, more informal, evidence bears this out. In the fourth year of the project, one student who had successfully completed both subjects wandered up to us in Week 4 and asked ‘How is this year’s mob progressing?’ In response to a non-committal answer, the student offered ‘They seem to be getting it. I overheard some of them having coffee this morning talking about what you are doing with them’ (RK). Of interest was the use of the preposition “with” and not “to” in this comment. Its use indicated to us the students thought of themselves as participants on a journey that was under their control and that the instructors were also participants in the learning and teaching that was occurring. This exemplified, for us, the desired outcomes of the participatory nature of the action research.

In short, although we had expected changes in students’ engagement, we were continually and pleasantly surprised by the extent to which task-orientated discussion happened within groups to the exclusion of other conversations, for example, their social lives. On one occasion, an important University official visited one practical class and talked with the students as they undertook their investigations. He was intrigued by the extremely high level of engagement, the buzz of interaction and the disparate nature of activity at each group’s table. Impressed by the professional manner in which the investigations were being conducted, he asked us ‘Is it always like this?’ an answer to which was given by one nearby student who interrupted with ‘Yes! Isn’t it exciting?’ (NM).

A high-stakes written assignment tendered for assessment in the final week of the semester draws upon both their early work in the subject and on their teaching and learning experiences that involved their explorations of various instructional strategies, which scaffolded their development of PCK. A final high-stakes criterion-referenced test is based on the same set of the ADT questions and the answers to which have to be supported by scientific explanations represents the final assessment and demonstration of their understanding of the content. The performance criterion in the test is set at 70% with scientific explanations having to be proffered to support their answer.

Within the assessment domain, both formative and summative feedback are provided by the instructors. In one major sense, the formal assignment responses elicited, and the feedback provided, by the instructors serve both a formative and summative role (Buck, Trauth-Nare & Kaftan 2010). Instructors provide feedback to the students on their performance in assignments through the criterion-referenced rubrics within four days of the attempt being submitted. The feedback is formative in the sense that if the standards have not been reached, students are allowed one resubmission in order to “make the grade” and summative in the sense that their first attempt is the mark that they “have to wear”. If a student who has not made the grade chooses to resubmit an acceptable attempt then they become eligible for a Passing Grade in the subject and marks are scaled in ways to reflect this achievement. This system means

they try their best on their first submission but they know that if it is not good enough, they will have an opportunity to address the feedback provided and re-submit.

The students also provide the authors with both forms of feedback through their responses to the one-minute feedback papers and through the formal summative subject evaluation process mandated by the university. All forms of feedback are considered and used to modify the subject offering in the following year. Evolutionary changes are explained to each new cohort of students together with the reasons for these modifications so that they understand the dynamic, interactive and evolving nature of learning and teaching within the subject design.

The Second Subject

The first subject serves as the platform upon which Subject 2 is built and which is undertaken in their fourth semester of study. Students have acquired significant skills in problem-based learning and cooperative learning strategies, and can apply these to address their lack of scientific content knowledge. Moreover, they understand the construct of PCK against which they can measure their development in terms of knowledge gains, facility with instructional strategies, knowledge of the curriculum, and of how students possess many explanations for their natural world, albeit not necessarily aligned with the scientific ones. These skills are developed further as they organise and jigsaw the many tasks described below, define the scope and sequence of content, identify appropriate activities, conduct various investigations to make sure that they work and, if not, to make adjustments, provide background knowledge/resources for any teacher who might use their unit, research the alternative conceptions that pupils might possess, define assessment procedures and feedback instruments and, communicate with each other effectively.

In Subject 2, we create a project-based learning context at the first meeting of a class where students assume the role of being “teachers” in an elementary school and the instructors their newly appointed Executive. The “Principal” invites the “Staff” to get into their “teaching teams or Stage Levels” (Kindergarten, Grades 1 & 2, 3 & 4, 5 & 6) in the same way as the mandated curriculum is organised. The instructors present a scenario constructed from the national and international research literature together with reports on the state of science education in Australia and internationally. The new executive ‘have examined your previous science programs and found them wanting’ in terms of the curriculum requirements of the Science Syllabus and of the amount of time devoted to the subject. The scenario is presented in such a way that they understand that the other science classes also represent “schools” in the same district each of whom will develop Units of Work for the other content strands of the elementary-school Science Syllabus and which gets drawn, at random, from a hat. Thus, collaboratively, the cohort will develop materials to cover all content areas.

In this context, the teaching teams brainstorm ideas and then apply a SWOT analysis (Strengths, Weaknesses, Opportunities, and Threats) before reporting the outcomes to the other teams within their class. Each group apportions tasks to its members for completion within a set time-frame as they begin to develop their Units of Work on their particular content area drawn from the Science Syllabus that articulates from Kindergarten through Grade 6. That is to say, the four units of work developed within a particular practical class maintain a “continuum of learning” from Kindergarten to Grade 6 in a particular content strand: Life Sciences; Physical Sciences; Earth and Space Sciences; and, Built Environments (the science content of the Australian Science Curriculum). These Units of Work have to involve school students in “Working Scientifically” and “Working Technologically” (the skills) as they undertake their investigations and which will also be designed to influence positively their attitudes and values. More importantly, for any investigation that is to be included, the group has to conduct it and provide feedback for the writer. In this way, they develop further their knowledge of content, instructional strategies, alternative conceptions and the curriculum documents. All investigations are built around Bybee’s (1997) 5Es instructional model. Influenced and motivated by the process, many students bring their laptops and tablets to supplement the six internet-connected computers in the classroom. Doing this also generates additional necessary information that is provided within their unit of work for those other teachers who may use it and who may not know it. They share their products using USB drives, email, Dropbox (or similar), closed Facebook groups that they set up and more informally through their many conversations and discussions within their own group, with their peers within their class, and also across the cohort.

At the staff meetings held at the start of a practical class each week, the instructors, acting as the Principal or Deputy, facilitate discussion as each group presents their latest outputs (the collaborative-learning products). At the end of the practical class each week, each group completes a meeting template, constructs an action plan and delegates tasks to be completed, individually or collectively, by the following week. These meeting minutes thus comprise a contract of work for the individuals within a group. The numerous roles are interchanged on a regular basis to ensure that all members acquire the skills of effective team management.

Halfway through the semester, their draft Units of Work are submitted for formative assessment with feedback provided within four days of submission. While acting on feedback, they also begin to make decisions about the activities that they will implement in local elementary schools, which we visit in the second half of the semester to conduct a science day. The purpose of this in-school experience is to develop further their existing PCK together with its additional two components of knowledge and beliefs about purpose, and knowledge of context.

In testing parts of their unit with real pupils, they observe the impact that it has both on the pupils and on the teachers who are required to be present and from whom feedback is elicited. Groups employ various mechanisms to obtain this feedback that is then used to fine tune the electronic hyperlinked version of their unit of work presented as a set of digital hyperlinked and interconnected documents, which is submitted for summative assessment at the conclusion of the semester. The in-school experience visit has a major impact on them as illustrated by a common comment such as this one ‘I enjoyed the school visit as it added meaning to the learning which allowed me to engage in the subject’ (SP). Indeed, when we investigated students’ written comments in 2011 with more frequent administration of the STEBI-B (eight occasions in 14 weeks), the impact of the in-school experience was reflected in a major increase in the PSTE score and which appears to be the largest increase in any one-week interval over the two subjects. In all other data sets that we collected from the Subject 2 cohorts, the STEBI-B was used only as a pre- and post-occasion instrument.

Treating our students as professionals in their field appears to be a major motivational factor that inspires them to act as teachers. Discussions center both on the impacts that their unit of work will have on pupils and on the coordination of various approaches, including curriculum integration, to improve the learning outcomes for school students. These high-level impact interactions are evidence of the professional way that they are being treated, their increasing confidence as teachers, and their developing expertise in programming and evaluation as illustrated by the following two pieces of feedback provided in the one-minute papers: ‘I feel the development of the unit was a rewarding and worthwhile exercise that allowed me to develop my skills as a unit developer and teacher’ (KD); ‘Overall, the subject layout was great – [It was] focused on our development as science teachers’ (KW). Summatively, many students comment positively on the scenario: ‘I liked the school/staff set up for classes. Thank you!!’ (SN); ‘Make sure you keep the “school” format etc. – It worked very well’ (ES). Moreover, they seem to understand the complementary nature of the two subjects as the following comment by a graduate illustrates:

In SC1 [the first subject] our lecturers [instructors] were teaching us the teaching styles required to teach science. ... In SC2 [the second subject] we were given the content and had to organise it into teaching experiences (GC).

An important additional motivational device introduced by us in the third year of the Project, 2008, involved us collating the many units of work and supplying them both to our students and to local schools on a CD-ROM. As the repository grew, the units of work were supplied on DVDs and more recently on 32 Gigabyte USB drives. It should be noted, however, that the Units of Work produced by the students are not components of the Content design element of Subject 2 in the Interactive Design Model. Instead, they are the outcomes of the processes embedded within the Interactive Design Model. Many students comment on the value of the collection:

The unit development assignment was fantastic and was a very worthwhile exercise. Very valuable resource provided on CD-ROM’ (CF).

Thus, our students exit this second subject with a full-school science teaching program covering all of the science content and skill components of the Elementary School Science Syllabus with parts that have been trialled and

evaluated. It also contributes to the professional learning of teachers in local schools, feedback from whom is extremely positive with many of them using the units of work with their classes.

It is also worth noting that there has been no duplication of the contents of any Unit of Work over a 10-year period. Interpretations of what can be done given the mandated outcomes are highly diverse. Thus, as the library of Units of Work grows, teachers can choose what is most appropriate for their class and make suitable adjustments to the digital materials to fit with their context. Evidence shows that the students are highly motivated, engaged and reflectively critical about their future roles with one student claiming:

This subject has been the most valuable core subject I have had so far – as we get useable units of work! There is something beneficial for us at the end of the subject and it makes the work worthwhile – not just to get a mark (MB).

We attribute these outcomes to the authentic project-based learning context into which they have been placed and the experiential learning in schools that has demonstrated to them that they *can* make a difference to the engagement and learning of pupils. Comments like the one above are not uncommon in the summative evaluations of the second subject provided by our students.

Instruments and Interactions

Of particular interest to us are the two constructs of *competence* and *confidence* (Appleton, 2003, 2007). We operationally define the construct of *competence to teach science* as the students' performance on the post-occasion Astronomy Diagnostic Test in the first subject. The Personal Science Teaching Efficacy belief scale of the Science Teaching Efficacy Belief Instrument B (STEBI-B) is used to define the construct of *confidence to teach science* (Enochs & Riggs 1990). The other scale in the STEBI-B is the Science Teaching Outcome Efficacy (STOE) belief. This is perhaps an unknowable construct to the preservice teachers given that they have no context on which to base the results of teachers' efforts other than their experience during their own schooling. Nonetheless, we were interested to track changes in their STOE scale scores to investigate if our students were potentially generalising their instructors' efforts within the educational design to the work of other teachers.

The Science Teaching Outcome Expectancy (STOE) belief scale has been reported in the literature to be somewhat problematic (Avery & Meyer 2012). Consequently, we modified the STEBI-B instrument in two ways. We employed an additional seven items from the Self-Efficacy Teaching and Knowledge Instrument for Science Teachers (SETAKIST) developed by Roberts and Henson (2000). Specifically, items 2, 7, 8, 13, 14, 15 and 16 of the SETAKIST were included in an attempt to bolster the construct of *confidence to teach science*. These additional items are:

- I do not feel I have the necessary skills to teach science in primary school.
- After I have taught a science concept once, I will feel more confident teaching it again.
- I find science a difficult topic to teach.
- I understand science concepts well enough to teach primary school science effectively.
- I know how to make primary school pupils interested in science.
- I feel anxious when teaching science content in primary school that I have not taught before.
- I wish I had a better understanding of the science concepts I will teach.

We subsequently computed a Confirmatory Factor Analysis on the resulting 30-items. Given that there were only 10 items related to Science Teaching Outcome Efficacy, of which two were problematic, we employed a parsimonious approach through reliability analyses to reduce the number of items in both scales to eight. We did this by simultaneously maximizing the Cronbach's alpha while retaining the power in Tukey's Test of additivity as close as possible to 1.000 thus rendering the resulting scales additive. Reliability Analysis reveals that both scales possess high levels of internal consistency for the PSTE and STOE scales (0.91 and 0.79, respectively) on the post-occasion of testing. On the pre-occasion of testing, the STOE belief produced a much lower Cronbach's alpha of 0.59. This latter finding is not entirely surprising given that our students had not experienced what efforts are required by teachers to generate the learning outcomes.

Each year, the items in the STOE scale yield a Cronbach's alpha of between 0.65 and 0.70 on the pre-occasion, and between 0.79 and 0.82 on the post-occasion. The lower value of the STOE scale on the pre-occasion is understandable given that our beginning students cannot truly know what a teacher has to do to improve pupils' performance. The Personal Science Teaching Efficacy (PSTE) belief scale yields a consistently higher Cronbach's alpha of 0.84 to 0.86 on the pre-occasion and 0.89 to 0.91 on the post-occasion. That is to say, students respond in consistent ways about their "confidence" to teach science. In addition, as students interact with the subjects, the internal consistency of the STOE scale rises indicating that their responses become more consistent as time passes. The STEBI-B data were collected over an eight-year period as an illustration of students' growth in "confidence" (Personal Science Teaching Efficacy (PSTE) belief) as well as tracking their belief that particular teacher behaviours can lead to better outcomes for students: the Science Teaching Outcome Efficacy (STOE) belief.

When we first introduced this instrument, we hypothesized that Subject 1 was likely to lead to a major reduction in the PSTE-belief scale score especially since what students thought they knew had been challenged and that their reasoning had been shown to be based on alternative scientific conceptions. We had further hypothesized that the STOE-belief scale score would show greater and positive changes because of the educational design where they began to experience personal success in learning science through our efforts and the design of Subject 1. That is to say, they were in a class where their instructors were modelling for them many approaches that appeared to covary with their perceived increases in competence.

They might have also come to perceive that good educational design and deep approaches to learning could generate the outcomes for their future students. This may have given them a level of assurance to invest in their own efforts to learn the science and how to teach it. We were surprised to find that the first hypothesis was wrong. In fact, there were positive, albeit small, changes in the PSTE scores during Subject 1. These changes became large in Subject 2 and especially after their in-school experience. The STOE belief showed consistent and significant growth when they were involved in both subjects.

The interaction between these two constructs was inferred from our observations that once the students understood that they could *actually* learn the difficult area of science covered in depth in Subject 1 (viz., the astronomy content), albeit at an elementary-school level, it seemed to engender a sense of *confidence* that they could learn other elementary-school science content and that they might thus become *competent* in learning and teaching the science contained in *any* of the content areas of the Science Syllabus. These inferences were confirmed in a deep investigation conducted within Subject 1 in 2010 where students completed the STEBI-B almost on a weekly basis (10 times in 14 weeks) and through interviews conducted on a biweekly basis with a sample (6) drawn from across the different classes (Cross 2011). In 2011, the same cohort completed the STEBI-B a further eight times. This was the only cohort of students to complete the instrument multiple times. To reiterate, all other cohorts since 2007 completed the STEBI-B only at the commencement and end of each semester of their studies in science and science methods (four times over two years).

Their emerging confidence is strengthened within Subject 2 where the content areas are deliberately chosen so that they do not include any aspects of the Space Science (astronomy) covered in Subject 1. Evidence for this interaction between the STOE and the PSTE beliefs is perhaps present in the PSTE scores in Subject 2 when they are required to go into schools to teach aspects of their Unit of Work. Indeed, in the second deep investigation conducted during 2011 with the same 2010 cohort of students, a highly significant increase in the PSTE-belief score happened in a very brief period of time (one week) and which covaried with their in-school experience when they realised the impact that the teaching of their materials had had on the school students.

RESULTS

The results presented here have been collected over the 10 years that this Participatory Action Research project has been in operation. Table 1 presents the effect sizes (Cohen's *d*) of three scales constructed from the ADT administered at the start of Subject 1 and again at its conclusion. It should be noted, however, that the second administration of the ADT is slightly different to the first insofar as the concepts tested are exactly the same but the questions explore different aspects of the same phenomenon. The responses made by students requires them to engage with the

Elaboration phase of the scientific concept in the 5Es instructional model (Bybee 1997). For example, in the first administration of the ADT, a question might ask ‘If the Earth’s orbit were changed to be exactly circular, what effect would this change have on the seasons?’ while in the second administration that question would be ‘If the Earth’s axis of rotation was changed to be at right angles to its orbit around the Sun, what effect would this change have on the seasons?’ Marks are awarded for the answer plus a valid scientific reason for that answer. Other questions contain overlays that require “visualization”, for example, in the pretest one question asks ‘Draw a picture to show how the Moon, Earth and Sun move’ while in the posttest the question is reframed as ‘Draw a picture to show how the Moon, Earth and Sun move from the perspective that you are in a spaceship high above the Earth’s south pole.’

Thus, the pre- and post-ADTs are not strictly parallel. In essence, the post-version can be described as more difficult. Thus, the effect sizes reported in Table 1 are likely to be underestimates of the “true effect” if exactly the same content test had been used on both the pre- and post-occasions. We make no apology for this since the assessment frame demands that the students demonstrate their *competence* within this subject and which adheres to the Bologna Model for Higher Education. The large effect sizes illustrate the extent to which changes covaried with the students’ presence in the class. The negative effective sizes for Alternative Conceptions indicates that the number of these reduced significantly from the pre- to the post-occasion of testing.

Table 1. Effect sizes of the Content Knowledge, Alternative conceptions, Complexity of Scientific Explanation in the Astronomy Diagnostic Test –indicators of *Competence*

| Effect sizes in: | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 |
|----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Content Knowledge | 2.13 | 1.08 | 5.95 | 5.75 | 2.69 | 2.41 | 4.50 | 4.24 | 4.12 |
| Alternative conceptions | -0.73 | -0.96 | -1.96 | -2.50 | -1.87 | -2.38 | -2.48 | -1.94 | -1.58 |
| Complexity of Explanations | 1.29 | 0.78 | 5.08 | 5.15 | 3.53 | 1.97 | 2.17 | 2.43 | 2.37 |
| N for paired data | 81 | 89 | 99 | 95 | 72 | 108 | 40 | 73 | 47 |

In educational research, to achieve a change of 0.7 sigma is notable where an effect size of 0.7 – 0.8 in is normally described as “strong” (Burns, 2000). To achieve changes of 1, 2, 5 and 6 sigma, as evident in Table 1, is testimony to the robustness of the educational design of the subject. It must also be said, however, that when one starts from a lowly mean of approximately 1.5 on the pre-occasion of testing, the large to very large effect sizes in content knowledge are not unexpected.

Encouraged by the 2-sigma effect in the first year of the Participatory Action Research project, we were somewhat disappointed by the 1-sigma effect obtained in the second year as far as Content Knowledge was concerned. In analysing the projects that the groups had undertaken, we noted that they had chosen to avoid those that contained difficult science concepts (e.g., phases of the Moon and the seasons) and had instead chosen those that they perceived to be easy or those with which they felt comfortable (e.g., writing a myth about a constellation or a poem about some astronomical phenomenon). That is to say, they avoided those projects that dealt directly with their alternative conceptions. Many described these projects as “too hard” for them. Thus in the third year, we provided a map from the questions in the ADT to the projects that they *must* do if they got the answer “wrong” on the pretest. This action had a dramatic effect on their performance on the posttest version of the ADT together with a dramatic improvement in the quality of their scientific explanations for the various phenomena. The variability in the effect sizes from year to year we find more difficult to explain. It may be due to the varying interests of students, to varying ability, to differences amongst cohorts, or to something that we have not yet identified. The 2010 cohort, for example, we identified as the weakest for some time yet they still achieved at a high level as indicated by the effect sizes greater than 2.00 for content knowledge and a significant reduction in alternative conceptions, and an effect size of almost 2.00 for the scientific reasons given for their answers.

Initially, for the two STEBI scales, we had intended to compute a multivariate analysis of variance with repeated measures on the occasion of testing using cohort membership as the independent variable. Unfortunately, when we checked the mathematical assumptions, the Box’s-M statistic was highly significant. We decided not to transform the data given the variation across scales and cohorts. For each of the two scales, we proceeded to compute separate ANOVAs with repeated measures on the occasion of testing for each complete set of cohort data whilst also checking

that the mathematical assumptions were met. A modified Bonferroni adjustment was employed given that six separate ANOVAS were computed for each of the two scales (i.e., 12 computations) and given that the average inter-scale correlation of 0.45 to yield a p-value of 0.01275 below which any differences can be deemed to be significant. Tables 2 and 3 present the descriptive statistics with effect sizes and levels of significance for the PSTE and STOE scales, respectively. No data were collected from the 2006/7 cohort in Subject 1 and data for Subject 2 experienced by the 2011/12 cohort is missing. Statistically significant increases to the PSTE subscales with moderate to strong effect sizes were reported on all but two iterations of Subject 1 (2007 and 2009) where very small effect sizes were detected.

Table 2. PSTE descriptive statistics with effect sizes and significance

| PSTE Subject 1 | | | | | | |
|----------------|--------------|-------------|---------------|-------------|--------------------|------------|
| Cohort Year | Pre Occasion | | Post Occasion | | ES/Sig | N |
| | Mean | Std. Dev. | Mean | Std. Dev. | Cohen's d | |
| 2006-2007 | - | - | - | - | - | - |
| 2007-2008 | 26.15 | 4.46 | 26.79 | 4.85 | 0.14** | 99 |
| 2008-2009 | 26.12 | 4.41 | 29.15 | 4.42 | 0.69*** | 95 |
| 2009-2010 | 27.65 | 4.55 | 27.76 | 4.80 | 0.02 ^{ns} | 72 |
| 2010-2011 | 26.09 | 4.11 | 28.39 | 5.77 | 0.46** | 108 |
| 2011-2012 | 25.93 | 3.50 | 28.95 | 6.23 | 0.60*** | 40 |
| 2012-2013 | 26.06 | 3.56 | 29.22 | 3.76 | 0.86*** | 73 |
| 2013-2014 | 25.02 | 3.35 | 27.51 | 4.43 | 0.63*** | 47 |
| All | 26.24 | 4.13 | 28.28 | 4.94 | 0.45** | 534 |

* p = 0.01275, ** p < 0.00255, *** p < 0.001275

| PSTE Subject 2 | | | | | | | |
|----------------|---------------------|-------------|---------------|-------------|----------------|------------|--|
| Cohort Year | Pre Occasion | | Post Occasion | | ES/Sig | N | Overall Cohen's d (PSTE Subject 1 & 2) |
| | Mean | Std. Dev. | Mean | Std. Dev. | Cohen's d | | |
| 2006-2007 | 22.24 | 4.51 | 28.94 | 4.85 | 1.43*** | 62 | - |
| 2007-2008 | 26.29 | 4.81 | 29.88 | 4.42 | 0.77*** | 76 | 0.84*** |
| 2008-2009 | 27.51 | 4.42 | 29.95 | 4.80 | 0.53** | 55 | 0.83*** |
| 2009-2010 | 26.33 | 3.83 | 29.10 | 5.77 | 0.57*** | 42 | 0.28* |
| 2010-2011 | 27.07 | 4.66 | 29.94 | 6.23 | 0.52*** | 67 | 0.73*** |
| 2011-2012 | Incomplete Data Set | | | | | - | - |
| 2012-2013 | 26.51 | 3.95 | 30.57 | 3.95 | 1.03*** | 49 | 1.20*** |
| 2013-2014 | 27.25 | 3.82 | 30.38 | 2.71 | 0.95*** | 16 | 1.76*** |
| All | 25.99 | 4.73 | 29.76 | 3.99 | 0.87*** | 367 | 0.87*** |

* p = 0.01275, ** p < 0.00255, *** p < 0.001275

Table 3. STOE descriptive statistics with effect sizes and significance

| STOE Subject 1 | | | | | | |
|----------------|--------------|-----------|---------------|-----------|---------------------|-----|
| Cohort Year | Pre Occasion | | Post Occasion | | ES/Sig | N |
| | Mean | Std. Dev. | Mean | Std. Dev. | Cohen's d | |
| 2006-2007 | - | - | - | - | - | - |
| 2007-2008 | 27.63 | 3.77 | 30.26 | 3.84 | 0.69 ^{***} | 99 |
| 2008-2009 | 27.48 | 3.38 | 30.45 | 3.43 | 0.87 ^{***} | 95 |
| 2009-2010 | 29.13 | 3.27 | 30.94 | 3.56 | 0.53 ^{**} | 72 |
| 2010-2011 | 28.51 | 3.48 | 31.32 | 3.47 | 0.81 ^{***} | 108 |
| 2011-2012 | 29.95 | 2.91 | 31.95 | 3.61 | 0.61 ^{**} | 40 |
| 2012-2013 | 29.66 | 3.63 | 31.81 | 3.77 | 0.58 ^{**} | 73 |
| 2013-2014 | 29.98 | 2.97 | 32.85 | 3.4 | 0.90 ^{***} | 47 |
| All | 28.64 | 3.54 | 31.17 | 3.66 | 0.70 ^{***} | 534 |

* p = 0.01275, ** p < 0.00255, *** p < 0.001275

| STOE Subject 2 | | | | | | | |
|----------------|---------------------|-----------|---------------|-----------|--------------------|-----|---|
| Cohort Year | Pre Occasion | | Post Occasion | | ES/Sig | N | Overall Cohen's d (STOE Subject 1 & 2) |
| | Mean | Std. Dev. | Cohen's d | Std. Dev. | Cohen's d | | |
| 2006-2007 | 29.5 | 3.39 | 30.45 | 3.06 | 0.29 [*] | 62 | - |
| 2007-2008 | 29.57 | 3.05 | 30.74 | 3.58 | 0.57 ^{**} | 76 | 0.85 ^{***} |
| 2008-2009 | 30.78 | 3.29 | 31.26 | 3.28 | 0.14 [*] | 55 | 1.14 ^{***} |
| 2009-2010 | 30.29 | 2.87 | 31.36 | 3.15 | 0.36 [*] | 42 | 0.69 ^{**} |
| 2010-2011 | 29.46 | 3.81 | 31.96 | 3.53 | 0.68 ^{**} | 67 | 0.98 ^{***} |
| 2011-2012 | Incomplete Data Set | | | | | - | - |
| 2012-2013 | 30.90 | 3.47 | 32.71 | 3.94 | 0.49 [*] | 49 | 0.81 ^{***} |
| 2013-2014 | 31.81 | 2.83 | 32.13 | 2.53 | 0.12 ^{ns} | 16 | 0.78 ^{***} |
| All | 30.08 | 3.37 | 31.38 | 3.46 | 0.38 [*] | 367 | 0.78 ^{***} |

* p = 0.01275, ** p < 0.00255, *** p < 0.001275

We note that between the end of Subject 1 in their first semester and the start of their fourth semester during which Subject 2 runs, there was a significant decrease in both the PSTE and STOE scale scores. We were curious about this reduction and investigated it through informal interviews with students. One indicator seemed to be repeated. Very few students had seen or experienced *any* science teaching in the interim period when they had undertaken their professional experience in schools. Perhaps this is one reason for the fall in PSTE and STOE beliefs. Others may be at play, for example, the changes in belief may not be stable when students do not have access to the subjects described here. We are investigating this hypothesis through following our graduates after they have completed these subjects and also after they have become teachers. The indications are that the PSTE belief continues to grow as they experiment and trial units of work with pupils in schools during their practicum sessions. Indeed, the PSTE belief appears to increase beyond the STOE belief achieved within these subjects. In addition, it appears that these increases in PSTE belief continue beyond their university studies and into their teaching careers ((Deehan, Danaia & McKinnon, *in review*).

DISCUSSION

Successive Australian government reports and international research literature recognise the importance of science in the elementary-school curriculum. The generalist nature of teaching in the elementary school is, however, a problem for the teaching of science in Australian schools. Tytler and colleagues (2008) identify both supports and barriers to science, technology, engineering and mathematics engagement in the middle years of schooling (upper elementary and junior high). They also identify pedagogy as being one key element in engaging students. Appleton (2003) focuses on the need for competent and confident elementary-school teachers. Many elementary-school teachers perceive themselves to possess neither of these two attributes. We argue that preservice elementary-teacher education faces a major challenge to generate and to build the competence and confidence of our future teachers if they are to remain generalists who will teach elementary-school science as well as the other curriculum areas.

Recent Australian discussions have indicated a preference for producing specialist elementary-school science teachers and to locate one within each school. We regard this as problematic given the opportunities that science and technology present for integration with other curriculum areas, for example, with mathematics and English as well as the Creative Arts and Social Sciences. Certainly, the specialist high school science teachers have not experienced concomitant success in Australia as indicated by Year 8 and 9 performance in PISA and TIMSS data (Thomson et al., 2016). For elementary-school science, it could be argued that employing specialist teachers will further separate ‘school science’ from the ‘real world’ where scientific literacy is required as well as denying opportunities for curriculum integration. Indeed, specialist science teachers in the elementary school rely on understandings developed in English and Mathematics as well as other curriculum areas.

Bybee’s (2014) article thus comes at an appropriate time in our country when we have been encountering the same problems of falling, or static, achievements in both PISA and TIMSS performance in Years 4, 8 and 9 while trying to develop strategies not only to counter this but also encourage more young people to undertake studies in science, technology, engineering and mathematics. We agree with both Tytler (2007) and Bybee (2014) who claim that teacher educators have to re–tune, replace, reform or even re–imagine science education in light of the publication of the *Next Generation Science Standards* in 2013 in the United States of America. We argue that these standards will require integrated approaches to the message systems of curriculum, pedagogy and evaluation (Bernstein, 1971) rather than the fine tuning of the extant practice of Education Faculty who focus on how to teach specific and different science topics each week while the science content is taught by Science Faculty in isolation from the instructional methods needed to teach it.

In this paper, we have presented an educational design model that emerged from the interrogation of our practice and which involves the students in deep, intense learning experiences, through integrating disciplinary knowledge with scientific practices, investigations and activities. We have taken this approach one step further by integrating the instructional strategies of how to implement the content within educational settings and within the investigations and activities that are driven by that content. The evidence that we have gathered during our 10–year participatory action research project, and which we have presented in a summary fashion here, serves to support what Bybee (2014) is asking of science educators albeit at the elementary-school preservice teacher–education level, i.e., ‘the reform of science teacher education should begin with the innovations of NGSS’ (Bybee 2014, p. 219). It must be stated, however, that our design is not a retrofit to Bybee (2014). Rather, it is an outcome of a deep scholarly approach that has occurred within a participatory action research frame and driven by the authors’ desires to see more and better science taught in elementary schools.

In interrogating our practice, an explanatory model emerged comprising three key design elements that employ PCK as the mediating domain of students’ interactions with the elements in mutually supportive ways. At the same time, the ‘model in operation’ involves a number of other well established approaches such as cooperative learning strategies, problem–based learning, formative feedback mechanisms and backward mapping of the assessment items to the learning experiences.

Our approach in Subject 2 suggests that the model can work in any science–content domain. Consequently, we would argue that any science area could be used as the vehicle in Subject 1 to drive students’ interactions with the relevant content. Thus, the model could be used more widely within other science domains, for example, force and motion, chemistry, the living world or in environmental science. Currently, we are involved in a project that involves ecology and sustainability to test this claim. Certainly, what has to happen is to remove the teaching of the science from the domain controlled by a Faculty of Science who lack the pedagogical insights of how science should be implemented in elementary schools and to locate it within the sphere of influence of Faculties of Education. This move, of course, has important implications for extant practice in universities if Bybee’s (2014, p. 217) assertions about the educational shifts are to happen.

When we visit other universities, we find that Australian Elementary School Teacher Professional Standards defined by various national bodies coupled with accreditation processes drive an interpretation that Science Faculty *should* be responsible for teaching the science content. This is entirely problematic given Bybee’s (2014) analysis. Historically, in the education domain we have had the ‘how to teach A’ and ‘how to teach B’ approach described by Bybee (2014,

p. 219) which clearly does not work in developing the greater senses of competence and confidence described by Appleton (2003, 2007).

What we are suggesting is to narrow the science content domain whilst also implementing the “how to teach it” component that Faculties of Education have adopted so that the “new” subject is like a hybrid of going deep in both the science *and* education domains simultaneously. The argument here is twofold: “less is more” and, once the preservice teachers understand that they *can learn* the science it produces an enduring sense of competence, that is, *that they can learn the science*. In Australia, these arguments are seldom implemented in teacher education courses. Rather, more is often added due to political pressure without due critical analysis of how the new interacts with that which is already present. That is to say, the curriculum for elementary pre-service teachers has become “a mile wide and an inch deep” (AAAS, 1989). Moreover, we are currently investigating the endurance factor of the educational design by following students into their third and fourth years of university and beyond when they are employed as elementary-school teachers. Evidence shows that the effects reported here are maintained and, indeed, grow beyond the end of their university studies and into their teaching careers. That is to say, their Personal Science Teaching Efficacy continues to grow as they encounter new science topics and engage in the personal science learning necessary in order to teach it to their pupils.

One clear implication from the approach described here is that the Education Faculty will have to be staffed with members who are very good in at least *one* science discipline in order to implement the model presented above and be experts in the teaching of that domain. Alternatively, if *scientists* are to teach preservice elementary-school teachers they will have to be well versed in pedagogy and the use of effective instructional strategies at the elementary-school level as well as in the assessment and evaluation of what they implement. A final possibility lies in the formation of *collaborative teams of scientists and science education experts* to deliver the content using appropriate learning experiences to generate the high levels of competence and confidence required of our elementary-school teachers when they enter the profession and to arrest the declines in elementary-school science performance. The second author is employing this approach in schools by constructing collaborative teams of elementary and high school teachers where the former bring their expertise in pedagogy and the latter their science content knowledge to the venture. The indications, at this stage, appear highly positive with a cross-over effect of increased confidence in the elementary teachers with the scientific content and more attentiveness being paid by the high school teacher to using appropriate instructional strategies to teach the content.

ACKNOWLEDGEMENTS

The authors profoundly thank all of our students in the Elementary Course for the honest feedback that they provided over the 10 years of our Action Research Project. They understood that they would not necessarily benefit from providing such feedback but were nonetheless willing to be participants in the Project for which the Ethics Protocol Number is 2006/122 at Charles Sturt University. New South Wales, Australia.

AUTHOR BIOGRAPHIES

David McKinnon is a Professor of Education at Edith Cowan University at Joondalup in Western Australia. He established and ran the Charles Sturt University Remote Telescope until his departure from Charles Sturt University in July 2014. He has a profound interest in science education at all levels and especially involving astronomy, public outreach and teacher professional learning. He has been involved with the Las Cumbres Observatory Global Telescope Network and its predecessor Faulkes Telescopes UK since 2003 as well as undertaking research into the impact of using the telescopes on students and teachers. David has been recognised nationally for his outstanding contributions to learning and teaching in science education with an Australian Learning and Teaching Council (ALTC) Teaching Excellence Award. E-mail: d.mckinnon@ecu.edu.au (corresponding author).

Lena Danaia is an academic within the School of Teacher Education at Charles Sturt University. She specialises in the science curriculum area. Lena previously worked as a teacher in both Australia and in the United Kingdom. Lena’s research in science education spans primary, secondary and tertiary contexts and has involved her working on a number of projects with a range of stakeholders. Lena’s PhD, awarded in 2008, focussed on students’ perceptions of,

and performance in, secondary school science both before and after their involvement in a project that fostered investigative, inquiry-based approaches in science teaching and learning and which provided students and teachers with the opportunity to control a telescope in real-time over the Internet. Lena has been recognised nationally for her outstanding contributions to learning and teaching in science with an Australian Learning and Teaching Council (ALTC) Teaching Excellence Award.

James Deehan graduated with a first class Honors degree from the Bachelor of Education (Primary) course in 2012. Since graduating James has worked casually in primary schools in the Bathurst region, with a particular foci on science education, gifted and talented education and small school education. The subject of James' Honors thesis was how the second science subject design influenced the science teaching efficacy beliefs of a cohort of pre-service primary teachers. In 2013 James was offered an Australian Postgraduate award to complete a PhD at CSU Bathurst. In his PhD research, James aims to determine how the science subjects offered at CSU influence the science teaching efficacy beliefs and science teaching practices of early career primary teachers.

REFERENCES

- Abrams, E., Southerland, S.A., & Evans, C. (2008). An introduction to inquiry. In E. Abrams, S.A. Southerland, & P. Silva (Eds.), *Inquiry in the classroom: Realities and Opportunities*, 1-13, Greenwich, Ct: Information Age Publishing.
- American Association for the Advancement of Science (1989). *Science for all Americans: Summary, Project 2061*.
- 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. (2007). Elementary science teaching, in S.K. Abell, & N.G. Lederman (eds), *Handbook of research on science education*, NJ: Lawrence Erlbaum, Mahway, pp. 493–535.
- Avery, L.M., & 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.
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.
- Bellocchi, A., Ritchie, S. M., Tobin, K., King, D., Sandhu, M., & Henderson, S. (2014), Emotional climate and high quality learning experiences in science teacher education. *Journal of Research in Science Teaching*, 51(10), 1301–1325.
- Bernstein, B. (1971). On the classification and framing of educational knowledge. In M.F.D. Young (Ed.). *Knowledge and Control: New directions for the sociology of education*. London: Collier MacMillan, 47–69.
- Bloom, B. S. (1971). Mastery learning. In J.H. Block & P.W. Airasian (Eds.) *Mastery learning: Theory and practice*, pp. 47-63. Holt Rinehart and Winston: New York
- Buck, G. A., Trauth-Nare, A., & Kaftan, J. (2010), Making formative assessment discernable to preservice teachers of science. *Journal of Research in Science Teaching*, 47(4), 402–421.
- Bybee, R. (1997). *Achieving scientific literacy: From purposes to practices*. Westport, CT: Heinemann.
- Bybee, R. W. (2014). NGSS and the next generation of science teachers. *Journal of Science Teacher Education*, 25(2), 211–221.
- Cantrell, P. (2003). Traditional vs. retrospective pretests for measuring science teaching efficacy beliefs in preservice teachers. *School Science and Mathematics*, 103(4), 177-185.
- Cochran, K. F., King, R. A., & DeRuiter, J. A. (1991). Pedagogical content knowledge: A tentative model for teacher preparation. ED 340683
- Collaboration for Astronomy Education Research (CAER). (1999). *Astronomy Diagnostic Test* as cited by Hufnagel, B., Slater, T., Deming, G., Zeilik, M., Adrian, R. L., Brick, C., & Adams, J. (2000). Pre-course results from the Astronomy Diagnostic Test. *Publications of the Astronomical Society of Australia*, 17(2), 152-155 as retrieved 20th May, 2004, from <http://solar.physics.montana.edu/aae/adt/> and modified by authors for Southern Hemisphere 2002 by the University of Sydney, further modified 2004 by authors at Charles Sturt University, Bathurst, NSW, Australia.
- Cross, M. K. (2011). The journeys of preservice primary teachers in learning how to teach science: A longitudinal case study. Bachelor of Education (Honours) thesis at Charles Sturt University, Bathurst, NSW, Australia.
- Deehan (2016). *The Science Teaching Efficacy Belief Instruments (STEBI A and B): A comprehensive review of methods and findings from 25 years of science education*. Springer International Publishing: Switzerland.
- Duch, B. J., Groh, S. E., & Allen, D. E. (2001). *The power of problem-based learning: A practical "how to" for teaching undergraduate courses in any discipline*. Sterling, VA: Stylus Publishing, LLC.
- 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.
- Franke, M. L., Carpenter, T. P., & Levi, L. (2001). Capturing teachers' generative change: A follow-up study of professional development in mathematics. *American Educational Research Journal*, 38, 653–689.

- Goodrum, D., & Rennie, L. (2007). *Australian school science education: National action plan 2008–2012 – Volume 1*. Canberra: Commonwealth of Australia.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. Teachers College Press, Teachers College, Columbia University.
- Hickey, D. T., Taasobshirazi, G., & Cross, D. (2012). Assessment as learning: Enhancing discourse, understanding, and achievement in innovative science curricula. *Journal of Research in Science Teaching*, 49(10), 1240–1270.
- Hoffman, D. H. (1978). Assertion Training: A Tool for Teacher Educators. ED224781.
- Hofstein, A., Eilks, I., & Bybee, R. (2010). Societal issues and their importance for contemporary science education2. In I. Eilks & B. Ralle (Eds.), *Contemporary science education* (pp. 5–22). Shaker, Aachen: Germany
- Johnson, D. W. (1994). Cooperative learning in the classroom. Association for Supervision and Curriculum Development, 1250 N. Pitt St., Alexandria, VA 22314.
- Johnson, D. W., Johnson, R. T., & Smith, K. A. (1998). Cooperative learning returns to college: What evidence is there that it works? Change: *The Magazine of Higher Learning*, 30(4), 26–35.
- 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.
- McIntyre, A. (2007). *Participatory Action Research* (Vol. 52). Sage Publications.
- Niess, M. L. (2005). Preparing teachers to teach science and mathematics with technology: Developing a technology pedagogical content knowledge. *Teaching and Teacher Education*, 21(5), 509–523.
- Palmer, D. (2007). Practices and innovations in Australian science teacher education programs. *Journal of Science Education*, 38(2), 167–188.
- Saxe, G. B., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understandings of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, 4, 55–79.
- Settlage, J. (2007). Demythologizing science teacher education: Conquering the false ideal of open inquiry. *Journal of Science Teacher Education*, 18(4), 461–467.
- Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., & Ruiz-Primo, M. A. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, 21(4), 295–314.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21, 43–69.
- Thomson, S., Wernert, N., O'Grady, E., & Rodrigues, S. (2016). *TIIMS 2015: A first look at Australia's results*. Camberwell, Victoria: Australian Council for Educational Research.
- Tytler, R. (2007). Re-imagining science education: Engaging students in science for Australia's future. Camberwell, VIC: Australian Council for Educational Research (ACER). Accessed 6 July 2015 from <http://research.acer.edu.au/acer/3>
- Tytler, R., Osborne, J., Williams, G., Tytler, K., & Cripps Clark, J. (2008). Engagement in STEM across the Primary–Secondary school transition: Opening up pathways. Canberra, ACT: Australian Department of Education, Employment and Workplace Relations. Accessed 6 June 2015 from <http://www.industry.gov.au/skills/Resources/Documents/OpenPathinSciTechMathEnginPrimSecSchTrans.pdf>
- Utley, J., Moseley, C., & Bryant, R. (2005). Relationship between science and mathematics teaching efficacy of preservice elementary teachers. *School Science and Mathematics*, 105(2), 82–87.
- Van Damme, D. (2001). Quality issues in the internationalisation of higher education. *Higher Education*, 41(4), 415–441.

NOTES