2010-07-20

Investigating the Reliability and Construct Validity of a Measure of Preservice Teachers' Self-efficacy for TPACK

Nicolette Michelle Smith
Brigham Young University - Provo

INVESTIGATING THE RELIABILTY AND CONSTRUCT VALIDITY OF A MEASURE OF PRESERVICE TEACHERS' SELF-EFFICACY FOR TPACK

Nicolette Burgoyne

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Charles R. Graham
Richard E. West
Richard R Sudweeks

Department of Instructional Psychology and Technology

Brigham Young University

April 2010

Abstract

Investigating the Reliability and Construct Validity

of a Measure of Preservice Teachers' Self-efficacy for TPACK

Nicolette Burgoyne

Department of Instructional Psychology and Technology

Master of Science

The TPACK framework is becoming increasingly pervasive in teacher education. Researchers and practitioners have been seeking reliable and valid ways to measure the constructs associated with the TPACK framework. This study describes the results of both an item review and the reliability and construct validity investigation of the scores from an instrument measuring self-efficacy for the constructs in the TPACK framework. Content-matter experts and the literature were used in order to perform the item review, while both an exploratory and a confirmatory factor analysis were performed in order to assess construct validity. Cronbach's alpha and Raykov's rho were used to assess the reliability. While the reliability was high, the validity was weak. Specific changes to the instrument were suggested as a means of improving validity.

Keywords: TPACK, self-efficacy, assessment, construct validity, reliability, exploratory factor analysis, confirmatory factor analysis

Table of Contents

List of Tables

List of Figures

Chapter I: Introduction

In 1986 Lee Shulman proposed a model consisting of the various domains of teacher knowledge: subject-matter content knowledge, pedagogical content knowledge, and curricular knowledge. Pedagogical content knowledge (PCK) is an amalgam of content knowledge and pedagogical knowledge and refers to the interpretations and transformations made by teachers on subject matter knowledge for facilitating the learning of students. As teachers apply their understanding of content, pedagogy, and their knowledge of learners to how particular topics to be taught should be represented and adapted to learners' characteristics and abilities, they are demonstrating PCK. Shulman (1986) defined PCK as, "the most useful forms of [content] representation … the most powerful analogies, illustrations, examples, explanations, and demonstrations – in a word, the ways of representing and formulating the subject that makes it comprehensible to others" (p. 9).

Since that time, researchers have built on Shulman's work in an attempt to understand PCK better, while often focusing on a particular content domain, such as mathematics (e.g. Hill, Ball, & Schilling, 2008). Grossman (1990), for instance, suggested four components of PCK: (a) conceptions of purposes for teaching subject matter; (b) knowledge of students' understandings, conceptions, and misconceptions of particular topics in a subject matter; (c) curricular knowledge; and (d) knowledge of instructional strategies and representations for teaching particular topics. This articulation, although not as clear in practice as it is in theory, has been helpful to researchers attempting to understand, research, and measure PCK more effectively.

In general, there is no definition or conception of PCK which is universally accepted. Van Driel, Verloop, and de Vos (1998) summarized the conceptualizations of PCK by various authors. They stated that some theorists included subject matter in the definition of PCK, while others included some combination of representations and strategies, student learning and conceptions, general pedagogy, curriculum and media, context, or purposes. Nevertheless, it is understood that PCK is concerned with teaching particular topics and involves teachers' knowledge of topic-specific representations and knowledge of learners' conceptions and misconceptions.

In recent years, though, technology has become an increasingly pervasive influence in people's lives as well as within various disciplines. This is one reason why it is necessary to incorporate technology into the teaching of various content areas (such as mathematics, science, and language arts) in order to equip students both in their future careers and lives. Various teacher educators have explored this problem and found it to be a complex issue. Many researchers have ignored the impact of the particular content domain in which the technology is being implemented (e.g. Ertmer, Conklin, Lewandowski, Osika, & Wignall, 2003; Hare, Howard, & Pope, 2002; Vannatta & Beyerbach, 2000).

Building onto the PCK framework, Mishra and Koehler have created a framework to explain the knowledge that teachers need to integrate technology into their teaching of a particular content area (Koehler, Mishra, & Yahya, 2007; Mishra & Koehler, 2006). This framework explicitly acknowledges that effective pedagogical uses of technology are deeply influenced by the content domains in which they are situated. For example, the teacher knowledge required to effectively integrate technology in a science classroom may be very different from that required for a social studies classroom. According to the framework, a teacher who can effectively integrate technology into the teaching of a particular content domain possesses technological pedagogical and content knowledge (TPACK).

To date, only a few researchers have attempted to create an instrument that measures an individual's knowledge of TPACK and its component parts (Archambault & Crippen, 2009; Archambault & Oh-Young, 2009; Cox, Graham, Browne, & Sudweeks, in review; Mishra & Koehler, 2006; Schmidt, Baran, Thompson, Mishra, Koehler, & Shin, 2009a; Schmidt, Baran, Thompson, Mishra, Koehler, & Shin, 2009). In 2008, a self-efficacy questionnaire for TPACK was created at Brigham Young University with the purpose of assessing the confidence in using TPACK (and its constructs) among pre-service teachers. While several of the surveys cited previously aim to measure teacher knowledge, the questionnaire in this study attempts to measure the self-efficacy of preservice teachers for TPACK. The motivation for measuring self-efficacy is that it is not simply a measure of knowledge and skills; rather it is a measure of what the respondent believes he or she can do.

With the frequent use of this questionnaire, it is necessary to determine if it is measuring what it was created to measure and whether the obtained scores provide a reliable measure of these constructs. Thus, the purpose of this thesis is to assess whether the scores from the self-efficacy questionnaire for TPACK are reliable and whether the interpretations of these scores possess validity.

This study focuses on the following questions related to the psychometric properties of the self-efficacy questionnaire for TPACK. It asks three main questions:

1. What evidence provided by exploratory and confirmatory factor analysis is there that the interpretations of the scores from the self-efficacy questionnaire for TPACK possess construct validity?
    a. What evidence is there that the structure underlying the items in the instrument is uni- or multi-dimensional?
    b. What evidence is there that the interpretations of the scores possess convergent and discriminant validity?
2. To what extent do the scores from the self-efficacy questionnaire for TPACK produce a reliable measure of each of the TPACK constructs?
3. How well do the current items in the instrument represent the domain of items they were intended to represent?

Chapter II: Literature Review

Both Shulman's model and Mishra and Koehler's framework led to new conceptions of teaching and teacher assessment. In this chapter the TPACK framework is explained and studies devoted to the development and review of assessments of both PCK and TPACK are considered.

TPACK Framework

With more teachers using technology in the classroom, Koehler and Mishra built on the notion of PCK to include the construct of technological knowledge and created the TPACK framework. The technological pedagogical and content knowledge framework describes the knowledges necessary for teachers to acquire in order to integrate technology into their teaching effectively (Koehler & Mishra, 2008). More specifically, this framework describes the complex interaction between a teacher's knowledge of the content (CK), pedagogy (PK), and technology (TK). This complex interaction results in four additional knowledges: pedagogical content knowledge (PCK), technological content knowledge (TCK), technological pedagogical knowledge (TPK), and technological pedagogical and content knowledge (TPACK), as shown in Figure 1.

Technological knowledge. Technological knowledge (TK) is the knowledge required to understand and use various technologies. These technologies may include both hardware and software. Basic TK might include simply an awareness that particular tools exist. More advanced TK, however, might include knowing how to use particular software programs or how to program in a particular language (Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b).

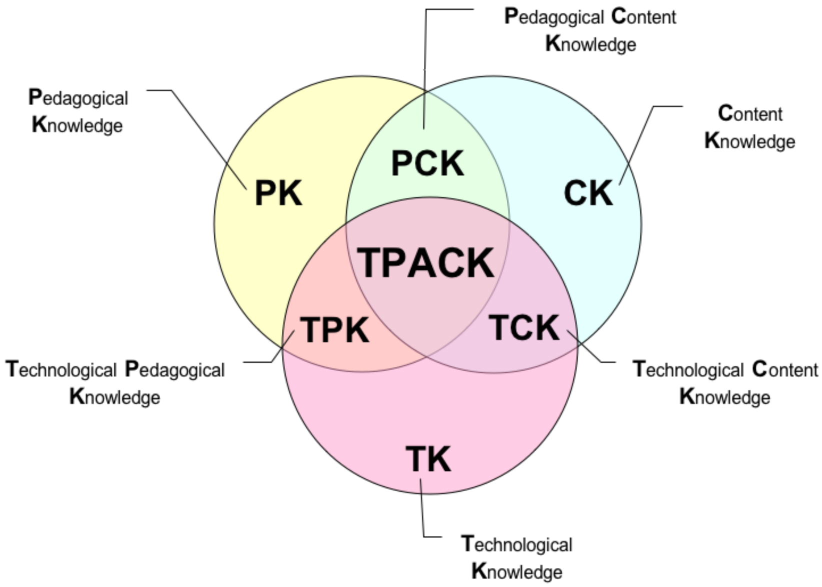

Figure 1. Visual representation of the TPACK framework.

Pedagogical knowledge. Pedagogical knowledge (PK) is the knowledge of the general processes and methods involved in teaching and learning multiple topics across multiple content domains. This may include knowing how to manage a classroom, motivate students to learn, as well as knowing how students learn, the developmental levels of students, develop and implement a lesson plan, assess students, or general teaching methods, such as discovery learning or collaborative learning (Cox, 2008; Harris, Mishra, & Koehler, 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b).

Content knowledge. Content knowledge (CK) is the knowledge of the facts, concepts, and skills of a particular content domain. This will include the methods for developing new knowledge as well as the representations of knowledge in that field (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006).

Technological pedagogical knowledge. Technological pedagogical knowledge (TPK) is the knowledge of how technologies can be used in a general (non-content specific) teaching context. This may include an understanding of how technology can be used to support teaching strategies and methods that can be used in any content area. For example, TPK may include knowing the basic rules for how to present information clearly using presentation software like MS PowerPoint, knowing when and how to use multimedia to engage an audience, knowing the strengths and limitations of online technologies for facilitating collaborative learning activities, and knowing what digital technologies and activities are appropriate for a particular age group (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006).

Pedagogical content knowledge. Pedagogical content knowledge (PCK) is knowledge of how to teach a particular content area. It is the knowledge of the analogies, illustrations, examples, explanations, and demonstrations that are effective in that content domain. It also includes a knowledge of common misconceptions or mistakes that students make as they learn that particular content, an awareness of students' prior knowledge, and a knowledge of content-specific pedagogies (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006). Thus, while pedagogical knowledge is the knowledge of how to teach using general pedagogical activities, Magnusson, Krajcik, and Borko (1999) stated that

> pedagogical content knowledge is a teacher's understanding of how to help students understand specific subject matter. It includes knowledge of how particular subject matter topics, problems, and issues can be organized, represented, and adapted to the diverse interests and abilities of learners, and then presented for instruction. (p. 96)

Technological content knowledge. Technological content knowledge (TCK) is a knowledge of the technologies that are relevant to a particular domain and how to use those technologies within the domain. TCK may include, for example, knowing how to use scanning electron microscopes to analyze insects. Additionally, TCK includes a knowledge of technology-enabled topic-specific representations used in the field (Cox, 2008; Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b).

Technological pedagogical and content knowledge. Technological pedagogical content knowledge (TPACK) is the knowledge of how to use technology to support content-specific pedagogical methods and strategies (or PCK) (Koehler & Mishra, 2008). There are two types of technological tools that might be used to support these content-specific methods: (a) content-domain oriented tools and (b) pedagogy oriented tools. Content-domain oriented tools are those technological tools learners may use that were created by practitioners in the particular content domain; for example, using data collection probes or measurement tools that a scientist might use in a scientific investigation. Pedagogy-oriented tools are those technological tools learners use that were created for a pedagogical purpose; for example, using a concept mapping tool, such as Kidspiration, that helps learners to visually organize information as they learn particular content (McCrory, 2008).

TPACK also involves the development of context-specific strategies and representations and how to coordinate these using emerging technologies in order to facilitate learning (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006). This includes an understanding of what makes certain concepts difficult or easy to learn and how technology can enable learning through the representation of these concepts.

Transformative versus integrative models. Gess-Newsome (2002), in speaking about PCK, suggested that one can consider a continuum of models of teacher knowledge. At one end of the continuum, there is the integrative model, where PCK is simply the intersection of content knowledge and pedagogical knowledge. At the other end of the continuum, there is the transformative model. In this model PCK is a new knowledge, where content knowledge and pedagogical knowledge combine into a unique form. Gess-Newsome compared these two models to chemistry. When two materials are combined, either a mixture or a compound can be formed. In a mixture, similar to the integrative model, the original elements remain distinct, though they may seem like a complete integration. In a compound, similar to the transformative model, the original elements cannot be separated nor their original properties identified.

Similar to these conceptions of PCK, some of the other constructs (in particular, TPK, TCK, and TPACK) can also be thought of in these ways. For example, using the integrative model, a teacher who possesses TK and PK would automatically also possess TPK, and a teacher who possesses TPACK simply possesses TK and PCK. However, if one uses the transformative model, a teacher who possesses TK and PK does not necessarily also possess TPK, since TPK is more than simply having TK and PK.

The model which one believes more closely resembles the relationship between these constructs will impact the nature of the items one constructs for an assessment. An integrative model would suggest that by combining aspects of TK and PK items, one can create TPK items. On the other hand, because a transformative model implies that a TPK item would be measuring a knowledge unique from the simple combination of TK and PK, TPK items would be completely distinct from TK and PK items.

Angeli and Valanides (2008, 2009) also related these models to the TPACK framework. They argued that the TPACK construct is a distinct body of knowledge and is constructed from a dynamic interaction between CK, PK, TK, and context; that is, they propose a transformative view of TPACK and they declare that they reject the integrative model.

Measurement of PCK

Since the TPACK framework incorporates PCK, the development of assessments measuring PCK is an important consideration when exploring how to assess TPACK. This section is kept brief since the focus of this thesis is on the measurement of TPACK, rather than PCK alone.

Kagan (1990, cited in Baxter & Lederman, 1999) stated that the challenge in assessing PCK is that it cannot be directly observed since it is partly an internal concept. Consequently, one cannot rely on observational data, since it provides only a limited view of a teacher's PCK, in that observers are not able to see the examples that the teacher does not use. For this reason, researchers have typically used self-report tests to gain an understanding of teacher's PCK.

Renfrow and Kromrey (1990) performed a review of research relating to assessments of teacher PCK. They provided some concrete examples of multiple-choice items used in the assessment of PCK, CK, and PK. The content-specific items that tested a teacher's PCK covered four main categories: (a) error diagnosis; (b) communicating with the learner; (c) organization of instruction; and (d) learner characteristics.

More recently Hill, Ball, and Schilling (2008) developed a measure of teachers' mathematical PCK. Their items fell into four categories: (a) common student errors; (b) students' understanding of content; (c) student developmental sequences; and (d) common student computational strategies. They used factor analysis, item response theory, and cognitive interviews to show multidimensionality of the item set as well as convergent and discriminant validity of the score interpretations. They found that the development of this instrument was challenging due to the underconceptualization of the constructs PK, CK, and PCK.

Measurement of TPACK

With the development of the TPACK framework, it became increasingly important to develop ways of measuring whether a teacher has TPACK (and its component parts) and is able to use this knowledge in practice. However, Archambault and Crippen (2009) have stated that TPACK is a difficult construct to measure because the seven parts of the framework seem confounded. Additionally, like the measurement of PCK, the development of assessments to measure TPACK is equally challenging due to the lack of consensus regarding the definitions of each of the constructs in the framework.

Mishra and Koehler (2006) were the first to develop a survey to measure TPACK, consisting of 33 Likert items and two short-answer questions. This survey, aimed at determining the level of TPACK knowledge both at the individual and the group level, was completed twice (at both the beginning and the end of the semester) by four faculty members and thirteen students. They found that the participants moved from viewing content, pedagogy, and technology as independent constructs towards a more unified understanding that indicated their development of TPACK.

Others have also used a pretest–posttest design (e.g. Schmidt et al., 2009a; Shin, Koehler, Mishra, Schmidt, Baran, & Thompson, 2009). Schmidt et al. (2009a) created a 50-item survey, where three of the items were open-ended questions asking for the respondent to describe specific situations in which TPACK was modeled and the remaining 47 items consisted of statements along with a 5-point Likert scale. Twelve of these items measured CK, seven measured TK, seven measured PK, four measured PCK, four measured TCK, five measured TPK, and eight measured TPACK. Eighty-seven preservice teachers enrolled in an introductory instructional technology class were asked to rate their knowledge. These preservice teachers also showed significant growth in all seven areas of the TPACK framework, but with the largest growth being in their TK, TCK, and TPACK. They also showed that the survey has an internal consistency (using Cronbach's alpha) between .75 and .92 for each of the seven constructs.

Using the same 50-item survey as described by Schmidt et al. (2009a), Shin et al. (2009) tested 23 graduate students also with the intention of determining how their understanding of the relationships between technology, content, and pedagogy changed over the semester. The results showed that the internal consistency (using Cronbach's alpha) for each sub-scale ranged from .40 to .98. They also showed that while the graduate students' TK improved, their CK and PK did not improve in general. In addition, their TCK, TPK, PCK, and TPACK improved.

On the other hand, other teacher educators have performed studies measuring teachers' knowledge in a particular instance, rather than examining growth over time.

Archambault and Crippen (2009) developed a survey consisting of 24 statements to measure teachers' knowledge. A national sample of 596 K–12 online teachers were asked to rate their own knowledge using a 5-point Likert scale (1 = *poor* and 5 = *excellent*) in terms of content, pedagogy, technology, as well as the overlapping areas created by merging CK, TK, and PK. They had twelve questions measuring PK, CK, TK, and TCK (three for each construct) and twelve questions measuring PCK, TPK, and TPACK (four for each construct), making 24 questions in total. In this study they established the reliability of the instrument and found that the internal consistency (using Cronbach's alpha) of the survey ranged from .699 to .911 for each of the constructs.

Using the same survey (but in web-form) and sample as Archambault and Crippen (2009), Archambault and Oh-Young (2009) found that these teachers rated their knowledge at the highest levels for PK (4.04), CK (4.02), and PCK (4.04), but were not as confident in their knowledge relating to technology (TK level at 3.04). Additionally, they found that the teachers' technological knowledge when combined with content or pedagogy increased.

Few researchers have as yet sought to establish the validity of the interpretations of the scores from their instrument. Schmidt et al. (2009b), in developing their TPACK survey, performed a pilot study on 124 students. They found Cronbach's alpha and used exploratory factor analysis on each domain. Using the results, 28 items of the original 75 items were deleted. Following this elimination, they found the internal consistency to range from .72 to .95 for each of the domains. The items in each of the domains of the TPACK framework loaded onto one factor, providing evidence for construct validity. A limitation of their study was that they only used exploratory factor analysis on each of the constructs and did not perform the analysis on the entire set of items. This makes it impossible to tell if the item set would load onto seven factors.

Archambault and Crippen (2009) also sought to establish the validity of the interpretation of the scores from their instrument by performing a think-aloud pilot. Participants were asked to explain what they were thinking as they answered each question. The researchers made several changes to the instrument after this pilot study in order to increase the construct validity of the interpretations of the scores from the survey.

Measurement of Self-efficacy for TPACK

Swain (2006) argues that although it may be evident that preservice teachers possess knowledge relating to technology integration, many do not believe that technology integration is worthwhile. Preservice teachers who possess this knowledge will not necessarily integrate technology into their future classrooms. Thus, measuring the knowledge and skills of preservice teachers is not a sufficient measure of whether they actually will use their newly acquired knowledge and skills. Measurement of self-efficacy, on the other hand, is a powerful predictor of future behavior, success, and persistence (Bandura, 1977; Multon, Brown, & Lent, 1991). In fact, Bandura (1977) stated that the stronger the perceived self-efficacy, the greater the effort will be.

Bandura (1994) defined perceived self-efficacy as “people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives. Self-efficacy beliefs determine how people feel, think, motivate themselves and behave” (p. 71). Similarly, Schunk (1984, cited in Milbrath & Kinzie, 2000) defined it as “personal judgments of one’s capability to organize and implement actions in specific situations that may contain novel, unpredictable, and possible stressful features" (p. 375). Thus, self-efficacy may be a mediator between knowledge and behavior. It is an individual's perceived self-efficacy that will enable them to translate their knowledge into behavior.

Several researchers have created instruments to measure the self-efficacy of preservice teachers for technology integration, where technology integration is a construct that can be seen as a combination of TPK and TPACK (e.g. Browne, 2007; Milbrath & Kinzie, 2000; Wang, Ertmer, & Newby, 2004). Consequently, these instruments measure a preservice teacher's belief in his or her ability to teach effectively with technology. One such instrument is the Technology Integration Confidence Scale (TICS), developed by Browne (2007), which is used to track preservice teacher self-efficacy for technology integration. This instrument presents a task and respondents then rate their confidence in accomplishing it. The TICS has six response categories for each item: *not confident*, *slightly confident*, *somewhat confident*, *fairly confident*, *quite confident*, and *completely confident*.

Recently, however, a few instruments have attempted to measure self-efficacy for TPACK (e.g. Cox et al., in review; Lee & Tsai, 2010). Cox et al. (in review) constructed a survey designed to measure the self-efficacy of preservice teachers for TK (ten questions), TPK (five questions), TCK (four questions), and TPACK (five questions). This instrument was the predecessor of the self-efficacy questionnaire for TPACK. Respondents were asked to rate their confidence (*not confident*, *slightly confident*, *somewhat confident*, *fairly confident*, *quite confident*, and *completely confident*) in their current ability to perform certain tasks. The survey was administered to nearly 200 preservice teachers. They found that the survey had high internal consistency. Furthermore, an exploratory factor analysis was conducted in an effort to determine whether the constructs as measured by the questionnaire were distinct. However, this analysis showed that the items loaded on only two major factors.

More recently, Lee and Tsai (2010) performed a study to investigate the perceived self-efficacy of teachers for a construct they called technological pedagogical content knowledge – web (TPCK-W), which emphasizes integrating web technology in the classroom. They developed a questionnaire, the TPCK-W Survey, to explore teachers' attitudes towards and self-efficacy for TPCK-W. The sample consisted of 558 elementary school and high school teachers in Taiwan. They then used exploratory and confirmatory factor analysis to explore the validity and reliability. The reliability for each construct was high. The exploratory factor analysis (EFA) revealed that the WPK and WPCK items loaded on the same factor. The confirmatory factor analysis revealed sufficient fit of the data to the model provided by the EFA. However, there were several limitations. Firstly, the researchers did not check to see if the data was normally distributed. Secondly, although the researchers performed an EFA before performing a confirmatory factor analysis (CFA) in order to know the structure of the model, they used the same data in both analyses. By using the same data, it is not surprising that the results of the CFA were good since the EFA produced a model fitted to the data.

As mentioned, there have been numerous questionnaires designed to measure self-efficacy for technology integration, and currently researchers are in the process of developing surveys to measure an individual's knowledge and skill in using TPACK and its component parts. However, attempts to create a survey that directly measures the self-efficacy for TPACK and its component parts is still in its infancy. This thesis hopes to contribute to the research in this arena.

Chapter III: Research Design and Methods

Context

Each semester preservice teachers majoring in elementary and early childhood education at Brigham Young University enroll in a required introductory instructional technology class. The course has been designed to help those enrolled develop knowledge, skills, and dispositions related to the use of technology in order to aid them in becoming more effective teachers. The class aims to teach them how to integrate technology into all content areas in the K–6 classroom.

Early on in the semester the students are taught about the TPACK framework. They then use this framework as they integrate technology with content and pedagogy in three assignments. These assignments involve the students creating a digital story, constructing a virtual tour, and using a technology that will aid in the teaching of science. The preservice teachers decide how they will use the assignments to teach students a particular content area. In addition, during each semester the preservice teachers who majored in elementary education do a four-week practicum where they focus on teaching language arts while in the schools. Students are encouraged to use the TPACK framework as they integrate technology into their lesson plans.

Participants

Those enrolled in this course are predominantly female and are either juniors or seniors. Before entering the introductory technology integration class, they are required to pass a basic proficiency test, called the technology skills assessment, which assesses their range of technological proficiency. This basic skills mastery test primarily deals with word processing, spreadsheets, PowerPoint, and internet communications.

The data consists of the responses from three groups of preservice teachers enrolled in the instructional technology class: those enrolled during the fall semester in 2008, the winter semester in 2009, and the fall semester in 2009. A description of these participants is provided in Table 1.

Table 1
Participant Description

	Fall 2008	Winter 2009	Fall 2009	Total
Completed the questionnaire	142	82	162	386
Gave permission to use results	125	75	133	333
Elementary Education majors	103	62	109	272
Early Childhood Education majors	22	13	24	61
Male	4	2	4	10
Female	121	73	129	323

Instrument

During 2008 an initial instrument measuring TK, TCK, TPK, and TPACK was created by Cox et al. (in review). After an initial exploratory factor analysis (EFA), the instrument was modified and items testing PK, CK, and PCK were added. This instrument became the self-efficacy questionnaire for TPACK used in this study.

The data for this study were collected through the use of this instrument. The questionnaire consists of 36 items. The number of items for each construct as well as the item codes can be found in Table 2. The respondents were asked to rate their levels of confidence (*not confident at all*, *slightly confident*, *somewhat confident*, *fairly confident*, *quite confident*, *completely confident*) with statements regarding their abilities to complete particular tasks (e.g. “Create a class website, blog, or wiki,” or “Use technology to teach language arts using content-specific methods (like balanced literacy, etc)”). Tables 18 to 24 in the Appendix contain the items for each TPACK construct from the self-efficacy questionnaire for TPACK.

Table 2
Item Summary

Scale	Number of Items	Item Code
Technological Knowledge (TK)	6	TK1 – TK6
Pedagogical Knowledge (PK)	4	PK1 – PK4
Content Knowledge (CK)	3	CK1 – CK3
Technological Pedagogical Knowledge (TPK)	7	TPK1 – TPK7
Pedagogical Content Knowledge (PCK)	4	PCK1 – PCK4
Technological Content Knowledge (TCK)	4	TCK1 – TCK4
Technological Pedagogical and Content Knowledge (TPACK)	8	TPACK1 – TPACK8
Total	36	

Data Collection

The data used in this study was collected at the end of three semesters: Fall 2008, Winter 2009, and Fall 2009. While a total of 386 preservice teachers completed the survey during these semesters, only 333 gave permission for their results to be used. The questionnaire was created using Qualtrics and the link to the survey was given to the students using the course management system. The results were then downloaded into an Excel spreadsheet.

Data Analysis

The following data analyses were performed (a) assessing the construct validity of the interpretations of the scores, (b) measuring the reliability of the scores obtained from the self-efficacy questionnaire for TPACK, and (c) reviewing the items.

Construct validity assessment. The labels TPACK, TPK, TK, TCK, and so on refer to abstract ideas (or constructs) created to assist in explaining the types of knowledge teachers have. Cronbach (1984, p. 133) stated that “a construct is a way of construing—organizing—what has been observed.” As has been stated, several tests have been created in order to assess these constructs. However, just because the tests have been created does not mean that the scores from the tests are valid dependable measures of that construct. Construct validation is the process whereby evidence is collected in order to support or refute a claim that a particular test is valid and measures the construct that the test developer claims it measures.

Table 3
Data Collection and Analysis Procedures for each Research Question

Research Question	Data Collection Procedures	Data Analysis
RQ1 – construct validity	TPACK questionnaire (36 items, n=333) administered at end of Fall 2008, Winter 2009, and Fall 2009 semesters.	Exploratory factor analysis was used to provide evidence for uni- or multidimensionality, while confirmatory factor analysis was used to provide evidence for convergent and discriminant validity.
RQ2 – reliability	TPACK questionnaire (36 items, n=333) administered at end of Fall 2008, Winter 2009, and Fall 2009 semesters.	Raykov's rho and Cronbach's alpha coefficients were computed to assess the reliability of the scale scores.
RQ3 – item review	TPACK questionnaire (36 items)	Through researching the literature and consulting subject matter experts, it was determined whether the items in the questionnaire are representative of the content domain.

To address the question regarding evidence for the construct validity of the scores and their interpretations obtained from the self-efficacy questionnaire for TPACK, both an exploratory factor analysis (EFA) and a confirmatory factor analysis (CFA) were performed.

Two main models were used. The first contains all the constructs in the TPACK framework (full model), while the second contains only those items involving technology (partial model). The partial model was examined because it was known that the items in the questionnaire currently measuring CK, PK, and PCK are not representative of the domain of all possible items measuring these constructs, since the instrument was designed to focus on items involving technology. While other items were added, the designers did not intend for them to be a comprehensive representation of the constructs CK, PK, and PCK. That is, it is known that part of the instrument lacks construct validity. Performing analyses on the partial model alone will enable one to assess whether this aspect of the instrument (consisting of the items involving technology) possesses construct validity.

Both the partial and the full models are transformative, where each construct is a different and unique knowledge. This is due to the fact that only transformative models can be tested using CFA. Brown (2006) states that CFA is used for specifying the number of factors, how the various indicators are related to these factors, and the relationships among indicator errors. However, Structural Equation Modeling (SEM) is needed in order to specify how the factors are related to one another, such as in an integrative model. Therefore, all the constructs specified by the TPACK framework are first order factors. Figures 2 to 3 display the models used in the analyses performed.

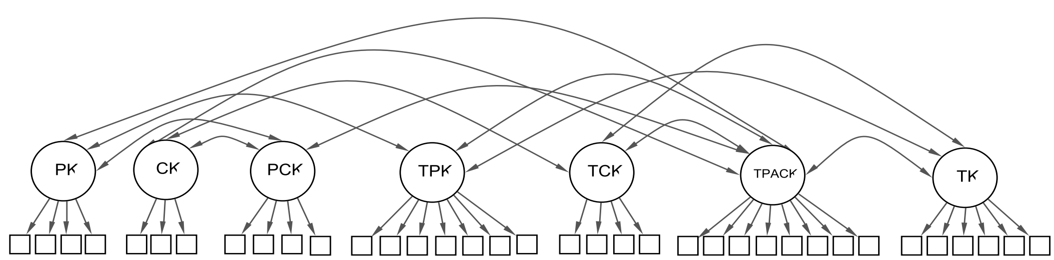

Figure 2. Path diagram of the full model.

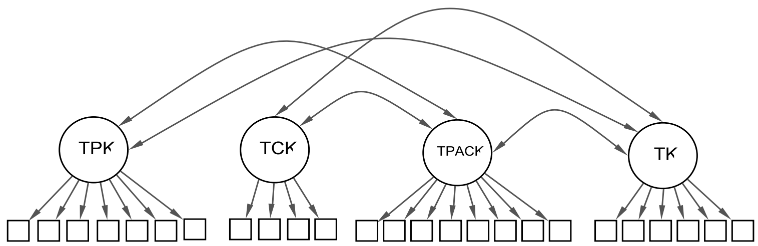

Figure 3. Path diagram of the partial model.

Two aspects of construct validity that were examined in this study are (a) the dimensionality of the item set and (b) whether there is convergent and discriminant validity.

Assessing the dimensionality of the items in the instrument enables one to know if there is a single underlying factor or if the underlying structure is multidimensional. If there is a single underlying factor, this would imply that one factor drives the responses to all the various aspects of the questionnaire and that the items which are supposed to be measuring different constructs are only measuring one construct. To explore whether the underlying structure is uni- or multidimensional an EFA was performed, since an EFA can show whether there is one factor that accounts for all the items in the questionnaire or whether there are multiple factors that account for these items as proposed by the theoretical TPACK framework.

Convergent validity implies that the items of a particular construct (i.e. TK or TPACK) should converge, which means that these items share a large proportion of variance (Hair, Black, Babin, Anderson, & Tatham, 2006). A CFA can be used to estimate the degree of convergent validity by examining the factor loadings and variance extracted. High factor loadings show that the items converge on some common point and that there is a greater amount of variance explained than error variance among each of the items. The variance extracted is another indicator of convergence, where it is hoped that each factor possesses more variance explained than error variance (Hair et al., 2006).

Discriminant validity, on the other hand, describes the extent to which a construct is distinct from other constructs. Using CFA, one can assess discriminant validity by comparing the variance extracted for any two constructs with the square of the correlation estimate between these two constructs. The square of the correlation coefficient represents the shared variance between the two factors. If there is discriminant validity, the two variance extracted estimates will be greater than the shared variance, since a factor should explain its items better than it explains another factor (Hair et al., 2006).

Reliability estimates. Reliability refers to the consistency of measurement and is generally determined by the overall proportion of true score variance to total observed variance of the measure. Cronbach's coefficient alpha (□) is often used to estimate reliability; however, often Cronbach's alpha is either an under- or an overestimate and therefore not dependable (Brown, 2006; Raykov, 2009). Raykov's reliability rho coefficient is found to be a more dependable estimate of reliability. Consequently, evidence for the reliability of the scores obtained from the self-efficacy questionnaire for TPACK was found using both Raykov's rho (using the results from a series of confirmatory factor analyses) and Cronbach's alpha. While Raykov's rho tests if a single factor underlies a set of variables (Raykov, 1998) and was calculated for each of the TPACK constructs, Cronbach's alpha was also calculated for each of the TPACK constructs.

Since the validity and reliability is built on evidence from multiple studies, the data from this study provides preliminary evidence towards establishing the validity and reliability of the instrument. Based on the results obtained, suggestions for improvement of the questionnaire were given.

Item review. In order to review the items in the questionnaire, the literature and five content-matter experts were consulted. These experts consisted of four professors outside of Brigham Young University who have specialized in TPACK research as well as one doctoral student who has done TPACK research. Each reviewer was asked via email to state whether they thought that the items were representative of each domain and what items were missing. Their feedback was then combined and based on this review, suggestions for improvement were provided. These suggestions include ideas for questions which ought to be asked based on aspects of the content domain that have not been assessed in the current questionnaire.

Chapter IV: Results

Construct Validity Assessment

In order to provide possible evidence for the construct validity, the data was first screened to determine if any problem items existed. Then an exploratory factor analysis was performed in order to provide evidence for multidimensionality of the item set. A confirmatory factor analysis was performed in order to determine model fit. The convergent and discriminant validity was examined using the results of the CFA.

Data screening. The normality of the data was evaluated by examining the skewness and kurtosis values for each of the items since the Maximum Likelihood estimation procedure used in the EFA and CFA analyses performed in this study assumes that the data follow a multivariate normal distribution. The distribution of responses to the items TK1 and TK2 both showed evidence of skewness and kurtosis. The skewness and kurtosis values of TK1 are -10.750 and 126.322 respectively and the skewness and kurtosis values of TK2 are -4.210 and 22.657 respectively. Both items have a skewness value that exceeds an absolute value of 2 and a kurtosis value that exceeds an absolute value of 7, which implies non-normality of the data (Finney & DiStephano, 2006). When examining the frequency distributions of both of these items it was evident that a majority of students (over 85%) felt completely confident in handling the tasks described in those items (sending an email with an attachment and using PowerPoint). For this reason these two items were removed and not included in the factor analyses.

Dimensionality. In order to assess whether the structure underlying the item set is uni- or multidimensional an EFA was performed using Mplus version 5.21. The Maximum Likelihood estimation procedure was employed because there are six categories in the rating scale. Since according to the framework many of the constructs are correlated with each other, the default oblique geomin rotation method of the factor pattern matrix was used.

The results for the EFA are shown in Table 4. Although a possible seven factors were specified, Mplus did not produce results for the six-factor model due to a lack of convergence; therefore, the results only show up to five possible factors. Using the Kaiser-Guttman rule of accepting factors with eigenvalues greater than 1.0 it seems that there are five factors since all the eigenvalues were greater than 1.0. By examining the ratios of the eigenvalues it is evident that there is one dominant factor and four less salient factors. The χ^2 statistic shows that the five-factor model is not a perfect fit but the other fit statistics show that it is the best fit of all the models.

Table 4

Summary of EFA Results when 1, 2, 3, 4, or 5 Factors are Extracted

Number of factors in model	Eigen-value	Eigen-value ratio[a]	Test of model misfit		Improvement in model fit			Summary Fit Statistics	
			χ^2	df	$\Delta\chi^2$	df	prob	CFI	RMSEA
1	15.694	6.93	3213.446	527				.693	.124
2	2.264	1.15	2420.220	494	793.226	33	0.00	.780	.108
3	1.973	1.15	1945.405	462	474.815	32	0.00	.831	.098
4	1.719	1.28	1609.426	431	335.979	31	0.00	.865	.091
5	1.338		1285.570	401	323.856	30	0.00	.899	.081

[a]Ratio of each eigenvalue divided by the next smaller one.

Table 5 displays the factor loadings of the 34 items on the five factors as well as the eigenvalues and percentage of variance for each factor. In order to increase meaningful interpretations of the results items having factor loadings of 0.40 and below were not reported in the results. Many of the items cross-loaded on multiple factors implying that the interpretation of the underlying factor structure is not clear-cut.

Table 5
Factor Loadings for Items in the Five-factor Exploratory Model

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
TK3	.526				
TK4	.875	.509			
TK5	.865	.536			
TK6	.599	.567			
PK1		.447	.737		
PK2			.702		
PK3		.522	.746		
PK4		.564	.733		
TPK1	.468	.818	.554		
TPK2	.521	.883	.401		
TPK3	.489	.757	.489	.433	
TPK4	.527	.838	.466	.404	
TPK5	.493	.881			
TPK6	.518	.876	.492		
TPK7	.518	.742	.402	.469	
PCK1		.523	.594	.612	
PCK2		.437	.594	.508	
PCK3		.450	.571	.774	
PCK4		.488	.521	.805	
TCK1	.589	.645			
TCK2	.587	.575			
TCK3	.537	.608		.714	
TCK4	.536	.688		.629	
CK1					.770
CK2					.743
CK3					.795
TPACK1	.437	.823	.483	.480	
TPACK2	.454	.816	.421	.473	
TPACK3	.488	.713		.512	
TPACK4	.448	.703	.406	.482	
TPACK5	.503	.710	.425	.714	
TPACK6	.406	.676		.560	
TPACK7	.468	.663	.412	.807	
TPACK8	.591	.696		.829	
Eigenvalue	15.694	2.264	1.973	1.719	1.338
% of Variance	68.270	9.850	8.580	7.480	5.820

The factor correlations are shown in Table 6. While none of the factor correlations are particularly high, factors 1 and 2 are the only pair that have a correlation exceeding .50. These factor loadings and factor correlations provide evidence for the lack of unidimensionality in the factor structure underlying the items and that there are multiple correlated factors underlying the item set.

Table 6
Factor Intercorrelations

	Factor 1	Factor 2	Factor 3	Factor 4
Factor 2	.585			
Factor 3	.231	.497		
Factor 4	.427	.484	.330	
Factor 5	.165	.276	.181	.233

Confirmatory factor analysis. The CFA was conducted using Mplus version 5.21. The raw data were used in each of the analyses. Since each of the items has six response categories the Maximum Likelihood estimator was used.

Using the χ^2 statistic and other fit indices, including the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and the comparative fit index (CFI), model fit was evaluated. Each index is important since each provides different information about model fit (Brown, 2006).

The χ^2 test is a measure of exact fit and a statistically significant χ^2 means that the model does not fit the data perfectly. However, χ^2 is sensitive to sample size and discrepancies to non-normality of the data; additionally it is limited because it is a test of exact fit (Byrne, 2005). This is another reason why additional fit indices (RMSEA, SRMR, & CFI) are helpful in assessing approximate fit. RMSEA tests for fit but adjusts for model parsimony and is therefore sensitive to the number of model parameters. SRMR, like χ^2, is an absolute fit index, testing for exact fit. CFI is an incremental fit index. These three fit indices range from 0 to 1. RMSEA and SRMR values closer to 0 and CFI values closer to 1 indicate adequate model fit.

Brown (2006) suggests that the following represent a reasonably good fit between the model and the data: RMSEA ≤ .05; SRMR ≤ .08; and CFI ≥ .95. However RMSEA values between .05 and .08 imply an adequate fit while values between .08 and .10 imply a mediocre fit. Values above .10 imply that the model should be rejected. Additionally CFI values between .90 and .95 indicate an adequate fit. It must be noted though that these criteria are not absolute cutoffs since there is no real consensus regarding what indicates a good fit.

In the models that were tested, the factor variances were constrained to 1 and all error covariances were set to 0. Using factor variances of 1 has the effect of standardizing the variance of each factor. The relationship between each pair of factors is then viewed as a correlation coefficient and is easier to interpret. Furthermore, constraining the factor variance ensures that Raykov's rho can be determined.

Model fit. Table 7 shows the degrees of freedom and the fit indices for both the tested models. Since the degrees of freedom for each model is positive, each of the models are over-identified. In both cases the $\chi 2$ test stated that the data is not an exact fit with the model ($p < .01$). Given that model fit cannot be based simply on the $\chi 2$ test, other fit indices were also calculated.

As was anticipated the RMSEA and SRMR values were close to 0 while the CFI value was close to 1. The RMSEA value for the full model was .080, which implies an adequate fit. Since the RMSEA value for the partial model was .088, which is greater than .080, but less than .100, the partial model is a mediocre fit. In both models the

SRMR values are less than .080 and thus a good fit is implied. In examining the value of the CFI the full model is a poor fit (since .878 is less than .90) while the partial model is an adequate fit with the data (since .908 is greater than .90). Thus the results are inconsistent. Brown (2006) states that when the fit indices provide inconsistent information about model fit, one needs to be cautious in deciding whether the solution is acceptable. The factor loadings and modification indices will provide possible reasons for this misspecification. This will then supply evidence regarding which items need to be changed or removed.

Table 7
Fit Statistics

Model	χ^2	df	RMSEA	SRMR	CFI
Full	1572.779	506	.080	.055	.878
Partial	798.253	224	.088	.053	.908

Note: RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index.

Parameter estimates. The standardized coefficients (□), standard error (s.e.), and variance explained (R2) were examined for both models (see Tables 8 & 9). All the coefficients in both models were found to be statistically significant at $p < .01$. Hair et al. (2006) suggest that standardized loading estimates should be at least .7 since a factor loading of .71 squared equals .5 (value of R2) which means there is more explained variance than error variance. However, loadings with values below .7 that are significant can still be considered but it should be remembered that there is more error variance than explained variance in these items. In the full model the standardized coefficients ranged from .536 to .888 while in the partial model the standardized coefficients ranged from .538 to .888.

Table 8
Standardized Parameter Estimates and Variance Explained (Full Model)

Construct	Items	λ	s.e.	R^2
Technological Knowledge	TK3	.536	.043	.287
	TK4	.868	.020	.754
	TK5	.865	.020	.748
	TK6	.632	.037	.400
Pedagogical Knowledge	PK1	.756	.029	.572
	PK2	.680	.035	.463
	PK3	.776	.027	.602
	PK4	.807	.025	.651
Content Knowledge	CK1	.745	.745	.555
	CK2	.740	.740	.548
	CK3	.825	.825	.681
Technological Pedagogical Knowledge	TPK1	.832	.018	.693
	TPK2	.884	.014	.782
	TPK3	.780	.023	.608
	TPK4	.848	.017	.719
	TPK5	.875	.015	.765
	TPK6	.888	.013	.789
	TPK7	.751	.025	.564
Pedagogical Content Knowledge	PCK1	.756	.027	.572
	PCK2	.680	.033	.462
	PCK3	.879	.017	.773
	PCK4	.871	.017	.759
Technological Content Knowledge	TCK1	.728	.029	.530
	TCK2	.721	.030	.519
	TCK3	.842	.020	.708
	TCK4	.845	.020	.715
Technological Pedagogical and Content Knowledge	TPACK1	.807	.021	.651
	TPACK2	.803	021	.645
	TPACK3	.746	.026	.556
	TPACK4	.730	.027	.533
	TPACK5	.831	.019	.690
	TPACK6	.744	.026	.554
	TPACK7	.817	.020	.667
	TPACK8	.847	.018	.718

Table 9

Standardized Parameter Estimates and Variance Explained (Partial Model)

Construct	Items	λ	s.e.	R^2
Technological Knowledge	TK3	.538	.043	.290
	TK4	.868	.020	.753
	TK5	.863	.020	.745
	TK6	.635	.037	.403
Technological Pedagogical Knowledge	TPK1	.829	.019	.687
	TPK2	.888	.013	.788
	TPK3	.780	.023	.608
	TPK4	.845	.017	.715
	TPK5	.876	.014	.768
	TPK6	.887	.013	.786
	TPK7	.752	.025	.566
Technological Content Knowledge	TCK1	.731	.029	.534
	TCK2	.726	.029	.527
	TCK3	.834	.021	.696
	TCK4	.846	.020	.716
Technological Pedagogical and Content Knowledge	TPACK1	.819	.020	.670
	TPACK2	.819	.020	.671
	TPACK3	.759	.025	.576
	TPACK4	.735	.027	.541
	TPACK5	.821	.020	.674
	TPACK6	.743	.026	.553
	TPACK7	.798	.022	.637
	TPACK8	.833	.019	.694

As mentioned earlier, squaring the standardized factor loadings produces each item's R^2, which is the variance in each item accounted for by the factor. In the full model the values of R^2 ranged from .287 to .789 (i.e. from 28.7% to 78.9% variance accounted for) while in the partial model the values of R^2 ranged from .290 to .788 (i.e. from 29.0% to 78.8% variance accounted for). Although most of the items had high variance accounted for by the respective factor, there were a few items with low variance accounted for by the factor and thus have large amounts of unexplained variance (or error variance). In the full model the items with at least 50% error variance are TK3 (71.3%), TK6 (60.0%), PK2 (53.7%) and PCK2 (53.8%), while in the partial model the items with at least 50% error variance are TK3 (71.0%) and TK6 (59.6%). This suggests that there is something about these items that is not explained by each of the models. Suggestions regarding the handling of these items will be given in the discussion section.

The variance extracted (VE) was also calculated for each factor. This is the average percentage of variance extracted among the items in each factor. Hair et al. (2006) suggest that a VE of 0.5 or lower suggests that on average there is more error in the items than variance explained by the factor structure imposed on the items. Table 10 shows the VE for each factor in both models. It is evident that for each of the other factors there is more explained variance than error variance.

Table 10
Variance Extracted for each Factor

Factor	Full Model	Partial Model
TK	.547	.548
PK	.572	---
CK	.594	---
TPK	.703	.703
PCK	.641	---
TCK	.618	.618
TPACK	.627	.627

Factor intercorrelations were also estimated (see Table 11). The theoretical framework for TPACK predicts that there should be a correlation between some of the constructs and no correlation between other constructs. The model states that there should be a correlation between TK and the other technology-related constructs (TCK, TPK & TPACK), between PK and the other pedagogy-related constructs (PCK, TPK, & TPACK) and between CK and the other content-related constructs (PCK, TCK, & TPACK).

However there should be no significant correlation between TK, PK and CK and minimal correlation between TPK, TCK and PCK.

Table 11
Factor Intercorrelations

Factor	TK	PK	CK	TPK	PCK	TCK
PK	.368**					
CK	.158*	.268**				
TPK	.641**	.652**	.309**			
PCK	.452**	.668**	.315**	.577**		
TCK	.691**	.542**	.337**	.768**	.700**	
TPACK	.635**	.646**	.313**	.872**	.774**	.858**

Note: *p<.05 **p<.01

The correlations between TK, PK and CK are low. However, the correlations between CK and all the other factors are low. This may be because there are only three items asking about CK confidence. Another possible reason for the low correlation may be that the CK items do not measure confidence of general content knowledge for a particular domain, such as science, but measure confidence of knowing one specific core curriculum statement in a particular domain. The correlations between TK and the other technologically-related factors (TCK, TPK & TPACK) are higher than the correlations between TK and the other non-technologically-related factors. Similarly, factor PK has higher correlations between the pedagogically-related factors (PCK, TPK and TPACK). Additionally there are large correlations between PCK, TPK and TCK. This implies that there is less differentiation between these items than desired and items would need to be altered for students to be able to answer these items differently, resulting in lower correlations.

The correlations between TPACK and TPK (r = .872) and TPACK and TCK (r = .858) are higher than desired. Brown (2006) states that correlations should be less than .85 in order for there to be differentiation between factors. Since these correlations are larger than .85 it is evident that participants may have struggled to differentiate between these items when completing the questionnaire. Another possible reason for the large correlations is the way in which the preservice teachers are being taught. This means that even if the items differentiate between the two constructs the preservice teachers are unable to because they have not been adequately taught the difference between these constructs.

Modification indices. Modification indices indicate approximately how much the $\chi 2$ statistic will decrease if the parameter (i.e. a factor loading or an error covariance), which was constrained before, is freely estimated. Implementing these changes may then improve the fit of the model. However, such changes need to be based on the theory. Since the theory explicitly states how the factors and items are to be related, modification indices rather suggest which items need to be altered or removed.

In the full model (see Table 12), in terms of cross-loading, it was suggested that TK6 cross-loads on the factors TPK, TCK, and TPACK. This item would probably need to be altered. Also, the items TPACK1, TPACK2, TPACK7, and TPACK8 cross-loaded on the factor TPK. This shows that there may be little differentiation between some of the TPK and TPACK items. Both TPACK2 and TPACK7 cross-loaded on PCK. In terms of item error covariances, there were large modification indices with PCK2 and TPACK6, PCK4 and TPACK8, TPACK1 and TPACK2, and TPACK7 and TPACK8. However, in the partial model (see Table 13) again TK6 cross-loaded on TPK, TCK, and TPACK. TPACK1 and TPACK8 cross-loaded on TPK. Regarding error covariances, there were large modification indices with TPACK1 and TPACK2; TPACK7 and TPACK8.

Table 12
Modification Indices (Full Model)

	M.I.
TPK by TK6	23.890
TCK by TK6	32.355
TPACK by TK6	30.319
TPK by TPACK1	39.387
TPK by TPACK2	29.860
TPK by TPACK7	24.698
TPK by TPACK8	31.960
PCK by TPACK2	24.886
PCK by TPACK7	27.560
PCK2 with TPACK6	57.749
PCK4 with TPACK8	63.677
TPACK1 with TPACK2	61.317
TPACK7 with TPACK8	98.038

Table 13
Modification Indices (Partial Model)

	M.I.
TPK by TK6	22.765
TCK by TK6	33.696
TPACK by TK6	33.106
TPK by TPACK1	29.536
TPK by TPACK8	22.586
TPACK1 with TPACK2	51.603
TPACK7 with TPACK8	116.097

Convergent validity. Hair et al. (2006) states that for there to be convergent validity the items of a construct should have a large proportion of variance in common. They suggest that the factor loadings and VE for each factor can be used in order to assess convergent validity.

In examining the factor loadings and their explained variance, the TK items did not appear to have a large proportion of variance in common in that TK3 and TK6 had factor loadings below 0.71 (and hence, more error variance than explained variance). Three out of four of the PK items and PCK items had a large proportion of variance in common while PK2 and PCK2 had more error variance than explained variance. Furthermore, the VE indicates that overall there is a large proportion of variance in common among each of the constructs, since each of the VE values are greater than .50. Consequently, overall, there seems to be an indication of evidence for convergent validity with certain items needing examination.

Discriminant validity. Hair et al. (2006) suggest that in order to assess the discriminant validity, the VE of each factor is compared with the shared variance between factors where the shared variance is calculated by squaring the correlation coefficients (found in Table 12). If the VE of two factors are both greater than the shared variance between these two factors, discriminant validity is then supported. For example, the values of the VE of TK and PK are .547 and .572 respectively, while the shared variance for TK and PK is .135. Since both of the variance extracted values are greater than the shared variance value, one can say that there is more different about these factors than is common, which supports the notion of discriminant validity. The values of the VE for each factor and the shared variance between each pair of factors are found in Tables 14 and 15.

Table 14
Comparison of Variance Extracted & Shared Variance of each Factor (Full Model)

Factor	VE	TK	PK	CK	TPK	PCK	TCK
TK	.547	---					
PK	.572	.135	---				
CK	.594	.024	.072	---			
TPK	.703	.411	.425	.095	---		
PCK	.641	.204	.446	.100	.333	---	
TCK	.618	.477	.294	.114	.590	.490	---
TPACK	.627	.403	.417	.098	.760	.599	.736

Table 15

Comparison of Variance Extracted & Shared Variance of each Factor (Partial Model)

Factor	VE	TK	TPK	TCK
TK	.548	---		
TPK	.703	.411	---	
TCK	.618	.480	.593	---
TPACK	.627	.404	.771	.734

While there is more different than is common among most of the factors in both models, it appears that there is more in common between the items testing TPK, TCK, and TPACK than is different. The lack of discriminant validity among certain factors (TPK, TCK, and TPACK) is confirmed by some of the modification indices, since there were several items that cross-loaded on these three factors. Thus, in order for there to be greater amounts of discrimination between these sets of items, changes will need to be made.

Reliability Estimates

In order to estimate the reliability of the scores of the self-efficacy questionnaire, Cronbach's coefficient alpha and Raykov's reliability rho coefficient were computed. Cronbach's alpha was computed using SPSS 16.0 while Raykov's rho was computed using parameter estimates obtained using Mplus version 5.21. Both of these coefficients were calculated to estimate the reliability of each of the subscales since these statistics assume that the set of items for which the statistic is calculated is unidimensional and it was shown that the entire scale is multidimensional. Thus, the reliability cannot be determined for the entire scale.

It must be noted that use of Cronbach's alpha makes two main assumptions. First, it assumes that the mean of the measurement error is zero (i.e. the error associated with the items are not correlated with each other); otherwise, Cronbach's alpha will be an overestimation of the true reliability. Second, it is assumed that the models of each construct are tau-equivalent. If this second assumption is violated, Cronbach's alpha will generally underestimate the reliability (Brown, 2006).

In both the models tested here, the covariances of the error terms for each item were each constrained to zero. Thus, the first assumption has been satisfied. In order to determine whether the second assumption is violated within each of the constructs it was necessary first to determine whether the model is congeneric. The congeneric model assumes that each item within a given scale measures the same construct or factor; that is, the congeneric model assumes that the set of items measuring a construct is unidimensional. The tau-equivalent model, on the other hand, is more restrictive since not only do the items measure the same construct but they measure it with the same amount of precision (although allowing for different amounts of error variance). Both the congeneric and the tau-equivalent models are nested and the fit of the data can be compared (using the fit statistics) from the least parsimonious (congeneric) to the most (tau-equivalent). However, tau-equivalence is rarely obtained and so Cronbach's alpha is often viewed as a lower bound estimate of reliability.

The fit indices of each model for each construct are shown in Table 16. In the case of the construct PK, the tau-equivalent model is the better fit (since $p>.01$) and regarding the construct TPACK, neither model is a good fit (since the RMSEA value and the minimum value of the 90% confidence interval is greater than .10 and the value of the CFI is less than .90). When the minimum of three indicators is used with no correlated errors, the model is just-identified and goodness-of-fit evaluation does not apply. Therefore, since there are only three indicators for CK, the model is just-identified and the goodness-of-fit indices in the CK congeneric model are zero. Thus, the fit indices for the CK congeneric model do not apply.

Table 16

Comparison of the Congeneric and the Tau-equivalent Models for each Factor

Model	χ^2	df	RM-SEA	90% CI	SR-MR	CFI	$\Delta\chi^2$	Δdf	p
TK – congeneric	16.735	2	.149	.088; .218	.028	.972	---	---	---
TK – tau-equivalent	310.526	5	.428	.389; .470	.345	.419	293.791	3	p<.01
PK – congeneric	9.176	2	.104	.043; .176	.020	.986	---	---	---
PK – tau-equivalent	14.108	5	.074	.030; .121	.091	.983	4.932	3	p>.01
CK – congeneric	0.000	0	.000	.000; .000	.000	.000	---	---	---
CK – tau-equivalent	11.278	2	.118	.058; .189	.132	.973	11.278	2	p<.01
TPK – congeneric	60.627	14	.100	.075; .126	.022	.976	---	---	---
TPK – tau-equivalent	80.486	20	.095	.074; .118	.094	.969	19.859	6	p<.01
PCK – congeneric	13.151	2	.129	.069; .200	.022	.984	---	---	---
PCK – tau-equivalent	115.726	5	.258	.218; .300	.345	.844	102.575	3	p<.01
TCK – congeneric	16.223	2	.146	.086; .216	.023	.978	---	---	---
TCK – tau-equivalent	41.650	5	.148	.109; .192	.133	.942	25.427	3	p<.01
TPACK – congeneric	246.931	20	.185	.164; .206	.049	.885	---	---	---
TPACK – tau-equivalent	317.571	27	.180	.162; .198	.147	.853	70.640	7	p<.01

Note: RMSEA = root mean square error of approximation; 90% CI = 90% confidence interval for RMSEA; SRMR = standardized root mean square residual; CFI = comparative fit index.

In each case, with the exception of the constructs PK, CK, and TPACK, the congeneric model is the better fitted model. The RMSEA values of the congeneric models for each factor are all greater than .10; however, the 90% confidence intervals have a minimum value less than .10. Additionally the SRMR values are all less than .08 and the CFI values are greater than .95. Therefore, these fit indices imply that the congeneric model in each case is a good fit. For this reason it would appear that the values of Cronbach's alpha are an underestimate of the true reliability of the items since the congeneric model is the better fitted model. In the case where the congeneric model is not a good fit (e.g. TPACK), implying a lack of unidimensionality, or where the congeneric model is just-identified (e.g. CK), Raykov's rho should be used cautiously.

Table 17 contains the values of Cronbach's alpha and Raykov's rho for each construct. The high reliabilities suggest that the items within each subscale are highly interrelated. This provides additional evidence for convergent validity.

Table 17
Reliability Coefficients for each Construct

Construct	Number of Items	Cronbach's Alpha	Raykov's Rho
TK	4	.792	.814
PK	4	.834	.829
CK	3	.811	.816
TPK	7	.939	.939
PCK	4	.871	.885
TCK	4	.862	.866
TPACK	8	.929	.932

Item Review

Five reviewers—all content-matter experts—were consulted and they reviewed the items found in the Appendix. Each reviewer was asked via email to state whether they thought that the items were representative of each domain and what items were missing. In general there was no consensus between the experts and some seemed to misunderstand the task required of them. It was hoped that the feedback obtained from the reviewers could be used in providing evidence for content validity. However, the data collected does not bear directly on this question and although useful, these results do not allow me to answer the question relating to content validity. Nevertheless, the feedback obtained has been combined and will be presented according to the constructs in the TPACK framework.

Technological knowledge. According to the literature, TK is the knowledge needed to understand and use various technologies. This covers a large range of skills, from simply knowing a technology exists to being able to program in a particular language (Koehler & Mishra, 2008). The current items in the TK section of the instrument focus on using a computer and therefore are probably too limited in their scope.

One reviewer suggested that the word “application” be used instead of “program” in TK5. Item TK6 was viewed as problematic by several reviewers, because it refers to the classroom and thus, it does not assess general TK. It should rather ask about confidence in creating a website, blog, or wiki. It was suggested that more items be added to this section in order to assess a larger range of technological skills. It was also recommended that the skills mentioned in the items extend beyond computer technologies and include digital cameras, cell phones, or iPods. Additionally, another question might address file management and where to find files that have been downloaded from the internet.

Pedagogical knowledge. According to the literature, PK consists of knowing the processes and methods involved in teaching. These processes and methods are general and not tied to a particular content domain. Therefore, classroom management strategies, motivational strategies, and so on fall into this type of knowledge (Harris et al., 2007; Koehler & Mishra, 2008). In the current instrument only four items are used to measure confidence in this knowledge type. This is a reflection on how limited this section is and how it is probably not representative of the domain of all possible items.

A comment was made about PK3 because it contains the word “lesson.” It was mentioned that lessons imply content, not just pedagogy, since a lesson always involves teaching specific content. It was suggested that this section include a larger range of skills, including understanding of instructional strategies and lesson planning. Another idea was to use the Interstate New Teacher Assessment and Support Consortium (INTASC) standards as a guideline for what items to include.

Content knowledge. According to the literature, CK consists of knowing the facts and concepts of a discipline and how they are typically represented, but also includes knowing methods for developing new knowledge within a particular domain (Harris et al., 2007; Koehler & Mishra, 2008). The items in the current survey do not reflect this definition.

No comments or suggestions were given regarding the items measuring this construct.

Technological pedagogical knowledge. According to the literature, TPK is knowing how technologies can be used in a general teaching context. This could be viewed as technology-enhanced PK (Harris et al., 2007; Koehler & Mishra, 2008). Although there are more items reflecting TPK skills than items that reflected PK skills, they probably do not cover the breadth of the domain of all possible items.

One reviewer suggested that the term "in the classroom" in TPK1 and TPK4 is unnecessary. It was suggested that TPK2 is too general and it may be better to ask about confidence pairing particular technologies with appropriate pedagogies. Another reviewer felt similarly about TPK2 and suggested that "technology-rich classroom" is too general a term and may be understood differently by different people. It was also questioned whether TPK5 is TPK or TK. Additionally, it was mentioned that the phrase "teaching productivity" may be interpreted differently by different preservice teachers. Either the phrase needs to change to something like "help me to plan and keep records to support my teaching" or an explanation in parentheses needs to be given. Finally, the comment was made that questions in the PK section can match the items in this section.

Pedagogical content knowledge. According to the literature, PCK consists of knowing how to teach a particular content area. Thus, the teaching strategies are no longer general but are content-specific. This includes knowing how best to represent the content in order to facilitate learning, being aware of the students' prior knowledge, as well as typical misconceptions and mistakes students make as they learn a particular content area (Harris et al., 2007; Koehler & Mishra, 2008). The current items are likely asking about specific skills in a way that is too general.

It was recommended that instructional strategies for each discipline be used to increase the number of items. One reviewer stated that the mathematics strategies currently listed are a bit weak and need to be clarified. A question regarding whether it would be appropriate to add items referring to content-specific pedagogies in the arts, physical education, and interdisciplinary topics was asked. Another suggestion was that questions should rather be framed along the lines of selecting the appropriate pedagogical strategy for particular content or learning goals. It was stated that PCK may be more about making connections between pedagogy and content rather than advocating a particular approach.

Technological content knowledge. According to the literature, TCK consists of knowing the technologies within a particular domain and the technology-enabled representations within a domain (Koehler & Mishra, 2008). The current items do not reflect this definition.

One reviewer noted that TCK4 is too vague and some specific tools for gathering data need to be specified. Another reviewer noticed that the words “use” and “used” are both in each item. It was recommended that the word “implement” or “integrate” be used to avoid the awkward construction. Other items may ask about finding and learning about content-specific technologies for each content area. It was suggested that an additional way to frame TCK is an understanding of how technology can change the content. For example, access to historical databases can expand the content of the curriculum. Furthermore, another reviewer stated that these items seem to measure TK, but are grouped by content areas. TCK may also be thought of as the knowledge teachers need to enable them to choose educational technologies appropriately in order to match the nature of particular content. Additional questions could be framed around this conceptualization of TCK. An item might then read, “Match technologies to mathematics content appropriately (e.g. Geometer’s Sketchpad).”

Technological pedagogical and content knowledge. According to the literature, TPACK consists of knowing the technologies that support content-specific pedagogies and topic-specific representations (Harris et al., 2007). These types of tools include pedagogy oriented tools and content-domain oriented tools (McCrory, 2008). TPACK can be thought of as technology-enhanced PCK. The current items are probably too general as they attempt to measure confidence in performing very specific skills.

There was a concern expressed that these items seem to be too parallel to the TPK items and it was suggested that another approach might be used where participants match the technology, instructional strategy, and content in a particular domain or perhaps assess the fit between content, technology, and pedagogy in a particular domain. In terms of particular item recommendations, one reviewer stated that TPACK1 is measuring TCK since no pedagogy is mentioned. TPACK2 and TPACK3 might be clearer if the word “curriculum” replaced “content.” Finally, TPACK4 might be clearer if it is framed in terms of supporting learning rather than supporting teaching. It was again noticed that the words “use” and “used” are in each item. It was recommended that the word “implement” or “integrate” be used to avoid the awkward construction.

It was also noted that an aspect that might be missing from the instrument are questions regarding context.

The interpretation and implication of these findings along with the other findings from this study will be discussed in the discussion section.

Chapter V: Discussion and Conclusion

The purpose of this research was to conduct both an item review and an initial investigation into evidence for the reliability and construct validity of the self-efficacy questionnaire for TPACK. The reliability was found to be satisfactory. Although some evidence for the validity was found, it is necessary to make several revisions to the questionnaire in order to increase the construct validity. The results of the factor analyses and the recommendations provided by the content-matter experts were used in making these suggestions for changes. These recommended item changes are discussed according to the construct they belong to in the TPACK framework and the ideas regarding what should be included in the questionnaire are found in Table 18.

Technological Knowledge

The first discovery was that items TK1 and TK2 describe tasks that are too easy for the participants completing the questionnaire. Since more than 85% of the participants feel confident sending an email with an attachment and creating a presentation using PowerPoint, it is strongly recommended that these items be removed.

In examining the modification indices of the CFA, TK6 cross-loaded with several factors (TCK, TPK & TPACK). Furthermore, the content-matter experts felt that this item was problematic since it refers to the classroom and therefore does not assess general TK. It is suggested that the word "classroom" be removed from this item in order to increase the validity.

According to the literature, items measuring TK should assess general technological skills (Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b). The content-matter experts suggested that these skills may extend beyond the domain of the computer. General technological skills that could be assessed in this section include the ability to use Microsoft office tools (i.e. Word, Excel, PowerPoint), the ability to use both a Mac and a Windows operating system, knowing various file formats (i.e. .doc, .pdf, .xls, .ppt, .kmz, .wmv, and so on), the ability to use the internet to find information (i.e. images, Wikipedia, etc), the ability to use a digital camera, the ability to use a digital microscope or digital probes (i.e. pH probe or temperature probe), and the knowledge of the affordances and constraints of various tools.

Pedagogical Knowledge

The results showed that PK2 was the only item that proved to be a problem in that it has more error variance than explained variance. This may be due to the notion of classroom management possessing different definitions in the minds of the participants, since it covers a fairly broad spectrum of skills. It is suggested that the several items referring to specific management strategies replace this single item asking about their confidence in classroom management in such a general way.

One reviewer noted that confidence regarding various general instructional strategies and lesson planning should be included. Another reviewer suggested that the INTASC be consulted when considering possible items for this section. One aspect of PK from these standards that has not been asked about is the ability to teach learners from diverse backgrounds.

Pedagogical knowledge includes knowledge of both general pedagogies and instructional strategies and of general learner characteristics. Thus, this section can be expanded to include items assessing knowledge and skill in both of these areas. For instance, items assessing general pedagogies may include specific classroom management strategies (efficiency, independent learning, pacing, student enjoyment, and so on), assessment strategies, collaboration strategies, motivational strategies, strategies that encourage hands-on learning, project-based learning, creating an authentic experience for the students, and strategies involving student practice of various skills. The section assessing general learner characteristics may include items assessing knowledge of various learning styles, knowledge of how to deal with students from diverse backgrounds, strategies that are age-appropriate, and the developmental characteristics of students (Cox, 2008; Harris, Mishra, & Koehler, 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b).

Content Knowledge

The EFA and CFA showed that these items converged and discriminated against the other items. Furthermore, the content-matter experts did not have any recommendations regarding them. However, the items were far too specific since they referred to a particular content standard and in examining the literature, CK is not only about knowing the facts and concepts within a particular content domain, but includes knowledge of the content representations in that field. Therefore, additional items may ask about knowledge of representations within each domain (e.g. the organization, editing, and publication of different genres in the language arts, the representation of data and calculations in mathematics, primary source documents and documentaries in social studies, and the collection and analysis of data in science). Additional items may also ask about confidence in using strategies for developing new knowledge and ways of thinking within each domain, such as scientific or historical ways of thinking (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006).

Technological Pedagogical Knowledge

Although there was little discrimination between the TPK and TPACK items, it is argued that this is due to the generality of the TPACK items. An example of this generality is in TPACK5: "Use technology to teach math using content-specific methods (like inquiry, standards-based math, etc)."

The content-matter experts did state however that TPK2 is too general. Similar to the PK item about classroom management, it is suggested that several items regarding various classroom management strategies be used.

The literature states that TPK is the knowledge of how technologies can be used in a general teaching context (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006). It is proposed that the TPK items be similar to the PK items (since both refer to a general teaching context), but paired with technology, as technology-enhanced PK. Thus, items can assess the confidence of participants in using general strategies with technology (i.e. technologies used for collaboration, technologies used for assessment, technologies used for motivation, and so on). Items can also assess confidence regarding knowledge of general learner characteristics and technology (i.e. age-appropriate use of technology and the prior knowledge of students in using various technologies).

Pedagogical Content Knowledge

It was shown that both PCK2 and PCK4 covaried with TPACK items (TPACK6 & TPACK8 respectively). PCK2 and TPACK6 are both about teaching language arts using content-specific strategies. Similarly PCK4 and TPACK8 are both about teaching science using content-specific strategies. The similarities between these items may have led to the preservice teachers answering them in similar ways. PCK2 also contains more error variance than explained variance.

The content-matter experts recommended that items ask about the participants' confidence in using domain-specific instructional strategies. One significant problem with the current PCK items is that they are attempting to ask about skills using specific strategies, but in a very general way. For example, PCK1 states, "Teach math using content-specific methods (like inquiry, standards-based math, etc)." These items need to ask about the skill of using content-specific methods in more precise ways, such as "Teach math using inquiry methods." Additionally, it was noted that there are no items regarding the arts or physical education.

According to the definitions provided in the literature, the construct PCK includes content-specific pedagogies, learner characteristics specific to content, and topic-specific representations relating both to pedagogy and the discipline, but transformed for teaching. Content-specific pedagogies can include both topic-specific strategies and content-specific strategies (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006). These elements are unique from the simple combination of CK and PK.

Although the current PCK items refer to these content-specific strategies, such as full scientific inquiry, balanced literacy, and so on, it is suggested that items ask the participants regarding their confidence in using these specific methods, rather than content-specific methods of science in general. Another aspect of PCK is learner content understanding. Items might ask about knowledge of students' prior knowledge and typical misconceptions in particular content areas. Finally, knowledge of representations used in teaching is part of PCK. These content representations may be pedagogical in nature, such as knowledge of using manipulatives in the teaching of mathematics, or the representations may be used in the discipline, but transformed for teaching, such as knowledge of using primary source documents in the teaching of social studies. It should be noted that while manipulatives and primary source documents are tools, they also provide a particular representation of the data which enables student learning.

Technological Content Knowledge

Although there was little discrimination between TCK and TPACK items, it will again be argued that this may be due to the generality of the TPACK items.

The reviewers suggested that TCK4 was too vague, but it is argued that each of these items is too vague. For example, TCK1 states, "Use technologies used in mathematics (like calculators, tools for creating graphs and charts, etc)." So the items should inquire about the confidence of using specific tools, rather than asking about the use of all technologies in math. For example, instead of asking about confidence in using technologies in mathematics, TCK1 could ask about confidence in using Excel, a tool used for creating graphs and charts in mathematics. Furthermore, TCK4 could ask about confidence in using digital probes, such as a temperature probe, which is used for collecting and analyzing data in science. It should be recognized that several items could inquire about the various tools used in each discipline.

The literature also states that TCK includes knowledge of the technologies in a particular field as well as technology-enabled topic-specific representations used in the field, which essentially is technology-enhanced CK (Cox, 2008; Koehler & Mishra, 2008; Mishra & Koehler, 2006; Schmidt et al., 2009b). Therefore, additional items may ask about the participants' confidence in using technologies that provide particular representations within each domain. An item might inquire about the confidence of knowing software used in the editing and publication of different genres in the language arts or of knowing a particular technology that enables graphical representations of data in mathematics.

Technological Pedagogical and Content Knowledge

This section was found to be problematic. The results of the modification indices showed that four of the eight items cross-loaded on TPK, while two items cross-loaded on PCK. In addition, the correlation between TPACK and TPK and the correlation between TPACK and TCK were too large showing little differentiation between the items measuring these constructs.

It is proposed that many of the items attempting to measure TPACK are too general and should be more specific. Describing specific strategies in a general way may have led to the high correlation between TPACK and TPK. Furthermore, the content-matter experts expressed their concern that these items are too similar to the TPK items.

One reviewer stated that TPACK1 is measuring TCK, but it is argued that this not valid since the item refers to improving student learning of a difficult concept. The mention of student learning means it is not a TCK item. It was suggested that TPACK2 and TPACK3 use the word curriculum instead of content. For example, TPACK2 would then read, "Use technology in a way that supports curriculum exploration and learning among your students." It was also suggested that TPACK4 refer to student learning instead of teaching. Thus, the item might read, "Identify online resources for students to use that support learning specific topics in the core curriculum."

Table 18

Suggested Ideas for the Items of each Construct

Construct	Ideas for Possible Items
TK	Knowledge of Microsoft office tools, the Mac and Windows operating systems, how to use the internet to find information, how to use a digital camera, how to use a digital microscope, how to use digital probes, various file formats, the affordances and constraints of various tools.
PK	Knowledge of how to encourage independent learning among students, how to pace the teaching of content, how to encourage student enjoyment of learning, various assessment strategies, collaboration strategies, how to motivate students, strategies that encourage hands-on learning, project-based learning, how to create an authentic experience for the students, strategies that encourage practice of particular skills, various learning styles, how to deal with students from diverse backgrounds, strategies that are age-appropriate, the developmental characteristics of students.
CK	Knowledge of the core content standards, the ways in which content is represented in each domain, strategies for developing new knowledge within each domain.
TPK	Knowledge of technologies used to encourage independent learning among students, technologies used to pace the teaching of content, technologies used to encourage student enjoyment of learning, technologies used for assessment, technologies used for collaboration, how to motivate students with technology, technologies that encourage hands-on learning, technologies that enable project-based learning, technologies that can create an authentic experience for the students, technologies that encourage practice of particular skills, how to use technologies to teach to various learning styles, how to use technologies with students from diverse backgrounds, technologies that are age-appropriate, the prior knowledge of students' use of particular technologies.
PCK	Knowledge of content-specific methods, for example, knowledge of full scientific inquiry methods, knowledge of balanced literacy methods, and so on, students' prior learning in math, science, language arts, and social studies, students' misconceptions in each domain, representations used in teaching each domain.
TCK	Knowledge of technologies used in the fields of math, science, language arts, and social studies, technology-enabled topic-specific representations used in the fields of math, science, language arts, and social studies.
TPACK	Knowledge of technology-enabled content-specific methods, for example, knowledge of technologies that encourage full scientific inquiry methods, knowledge of technologies that encourage balanced literacy methods, and so on, how technologies can help overcome students' misconceptions in math, science, language arts, and social studies, technologies that enable representations used in teaching math, science, language arts, and social studies.

As defined in the literature, TPACK is the knowledge of how to use technology to support content-specific teaching methods, an understanding of how technology can help learners overcome misunderstandings of particular concepts, and knowledge of how technologies can facilitate learning through the representation of concepts (Cox, 2008; Harris et al., 2007; Koehler & Mishra, 2008; Mishra & Koehler, 2006). In examining the definition of PCK, it is evident that TPACK can be viewed as technology-enhanced PCK. Thus, the TPACK items can map onto the PCK items, but specify particular technologies; that is, items regarding specific technologies that can be used with content-specific pedagogies, technologies that assist in learner understanding and overcoming the misunderstanding of content, and technologies that enable topic-specific representations of content can be used (i.e. using virtual manipulatives in mathematics, using specific online tools for writing and publishing in language arts, using mini-documentaries in social studies, using probes for collecting and analyzing data, etc).

Conclusion

In this study, both a full model (consisting of the entire item set) and a partial model (consisting of only those items relating to technology) were tested. The results for both models showed that regardless of whether all the items are used or only those relating to technology, the models were mediocre fits with the data. This lack of a good fit may be due to localized misfit as evidenced by the lack of discriminant and convergent validity of some constructs.

These models were transformative, since this is the only type of model that can be tested using CFA. However, in examining the definitions of the certain constructs provided by the literature (TPK, TCK, & TPACK) a transformative model does not seem to be an adequate way of describing the relationships between the constructs since TPK appears to be technology-enhanced PK, TCK appears to be technology-enhanced CK, and TPACK appears to be technology-enhanced PCK. Thus, an integrative model would appear to be a better way of representing the relationships between these constructs (Gess-Newsome, 2002). On the other hand, in looking at the definitions provided by the literature of PK, CK, and PCK, PCK is clearly a knowledge that is very different than PK and CK. Therefore, a transformative model is a better representation of the relationship between these constructs (Gess-Newsome, 2002). An alternative model (lying on the continuum between the transformative and integrative models) is suggested; this model may include the relationship between PK, CK, and PCK being viewed from a transformative model perspective (PCK is a new unique form when PK and CK are combined) and the constructs TPK, TCK, and TPACK being viewed from an integrative model perspective (TPK is the combination of TK and PK, TCK is the combination of TK and CK, and TPACK is the combination of TK and PCK). Instead, Angeli and Valanides (2008, 2009) suggested that the TPACK construct be viewed from a transformative perspective since they proposed that it results from the dynamic interaction between TK, PK, CK and the context one is teaching in. They argued this approach since they believed that a teacher who possesses TK, PK, and CK does not necessarily also possess TPACK.

It has been shown that the scores of the current instrument possess high amounts of reliability and that the interpretations of these scores possess some construct validity. It is believed that implementing the suggested changes to the items will increase this validity. As has been stated, these recommendations imply that an alternative model (lying between a transformative model and an integrative model) is the best way to represent the relationship between the constructs. However, given that CFA can only use transformative models in the analyses, alternative methods would need to be used to determine if this alternative model is a good fit with the data.

For future research it is recommended that substantial changes be made to the questionnaire according to the suggestions proposed and then, using results obtained from preservice teachers, structural equation modeling can be used to verify whether the alternative model put forward here is a good fit with the data.

A limitation of this study was the inability to provide evidence towards content validity and adequately answer the question regarding whether the items are representative of the possible domain. Future research can focus on this aspect of the instrument.

Another limitation of this study (and other possible future studies) is that there is no real consensus in the definitions of the constructs in the TPACK framework. This makes it difficult to state whether an instrument measuring TPACK or self-efficacy for TPACK really measures what it is supposed to measure. However, in spite of this limitation, it is hoped that not only will these recommendations increase the construct validity, but that these item recommendations can be considered for use by other TPACK test developers as necessary knowledge for teachers who integrate technology in their teaching.

References

Angeli, C. & Valanides, N. (2008). TPCK in pre-service teacher education: Preparing primary education students to teach with technology. Paper presented at the *2008 Annual Meeting of the American Educational Research Association*. March 24-28, New York City, New York.

Angeli, C. & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT-TPCK: Advances in technological pedagogical content knowledge (TPCK). *Computers and Education, 52*, 154-168.

Archambault, L., & Crippen, K. (2009). Examining TPACK among K-12 online distance educators in the United States. *Contemporary Issues in Technology and Teacher Education, 9(1)*. Retrieved March 15, 2009, from http://www.citejournal.org/vol9/iss1/general/article2.cfm

Archambault, L., & Oh-Young, C. (2009). Putting the T in PCK: Exploring the nature of the TPACK framework among K-12 online educators using a web-based survey. Paper presented at the *2009 International Conference of the Society for the Information and Technology and Teacher Education*. March 2-6, Charleston, South Carolina.

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

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. Retrieved April 7. 2010, from http://www.des.emory.edu/mfp/BanEncy.html

Baxter, J.A. & Lederman, N.G. (1999). Assessment and measurement of pedagogical content knowledge. In J. Gess-Newsome & N.G. Lederman (Eds.), *Examining pedagogical content knowledge*. (pp. 147-162). Norwell, MA: Kluwer Academic Publishers.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: The Guilford Press.

Browne, J. (2007). *Evidence supporting the validity on inferences required by the intended uses of the technology integration confidence scale*. Doctoral Dissertation, Brigham Young University, Provo, UT.

Byrne, B.M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment, 85(1)*, 17-32.

Cox, S. (2008). *A conceptual analysis of technological pedagogical content knowledge*. Doctoral Dissertation, Brigham Young University, Provo, UT.

Cox, S., Graham, C. R., Browne, J., & Sudweeks, R. (in review). How do we measure TPACK? An attempt to create and validate an instrument.

Cronbach, L.J. (1984). *Essentials of psychological testing* (4th ed.). New York: Harper & Row.

Ertmer, P., Conklin, D., Lewandowski, J., Osika, E., & Wignall, E. (2003). Increasing preservice teachers' capacity for technology integration through the use of electronic models. *Teacher Education Quarterly, 30(1)*, 95-112.

Finney, S.J. & DiStephano, C. (2006). Nonnormal and categorical data in structural equation models. In G.R. Hancock & R.O. Mueller (eds.). *A second course in structural equation modeling.* (pp. 269-314). Greenwich, CT: Information Age.

Gess-Newsome, J. (2002). Pedagogical content knowledge: An introduction and orientation. *In examining pedagogical content knowledge.* (pp. 3-17). Retrieved October 27, 2009, from http://dx.doi.org/10.1007/0-306-47217-1_1.

Grossman, P.L. (1990). *The making of a teacher: Teacher knowledge and teacher education.* New York: Teachers College Press.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L. (2006). *Multivariate data analysis*. New Jersey: Pearson Prentice Hall.

Hare, S., Howard, E., & Pope, M. (2002). Technology integration: Closing the gap between what preservice teachers are taught to do and what they can do. *Journal of Technology and Teacher Education, 10(2)*, 191-203.

Harris, J., Mishra, P., & Koehler, M.J., (2007, April). Teachers' technological pedagogical content knowledge: curriculum-based technology integration reframed. Paper presented at the *2007 Annual Meeting of the American Educational Research Association*, Chicago, IL.

Hill, H.C., Ball, D.L., & Schilling, S.G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39(4)*, 372-400.

Koehler, M.J., Mishra, P., & Yahya, K. (2007). Tracing the development of teacher knowledge in a design seminar: Integrating content, pedagogy, & technology. *Computers & Education, 49(3)*, 740-762.

Koehler, M. J., & Mishra, P. (2008). Introducing technological pedagogical content knowledge. In AACTE (Eds.), *The handbook of technological pedagogical content knowledge for educators*. (pp. 3-30). New York, NY: Routledge/Taylor & Francis Group for the American Association of Colleges of Teacher Education.

Lee, M. H. & Tsai, C. C. (2010). Exploring teachers' perceived self-efficacy and technological pedagogical content knowledge with respect to educational use of the World Wide Web. *Instructional Science, 38(1)*, 1-21. Retrieved March 19, 2010, from http://www.springerlink.com/content/d510480505435104/

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining Pedagogical Content Knowledge*. (pp. 95-132). Norwell, MA: Kluwer Academic Publishers.

McCrory, R. (2008). Science, technology, and teaching: The topic-specific challenges of TPCK in science. In AACTE (Eds.), *The handbook of technological pedagogical content knowledge for educators*. (pp. 193-206). New York, NY: Routledge/Taylor & Francis Group for the American Association of Colleges of Teacher Education.

Milbrath, Y. L. & Kinzie, M. B. (2000). Computer technology training for prospective teachers: Computer attitudes and perceived self-efficacy. *Journal of Technology and Teacher Education, 8(4)*, 373-396.

Mishra, P. & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A new framework for teacher knowledge. *Teachers College Record, 108(6)*, 1017-1054.

Multon, K., Brown, S., & Lent R. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38.

Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement, 22(4)*, 375-385.

Raykov, T. (2009). Evaluation of scale reliability for unidimensional measures using latent variable modeling. *Measurement and Evaluation in Counseling and Development*, *42(3)*, 223-232.

Refrow, D.D. & Kromrey, J.D. (1990). Descriptions and uses of content-specific pedagogical items. Paper presented at the *1990 Annual Meeting of the National Council on Measurement in Education*. April 17-19, Boston, MA.

Schmidt, D., Baran, E., Thompson, A., Mishra, P., Koehler, M.J., & Shin, T. (2009a). Examining preservice teachers' development of Technological Pedagogical Content Knowledge in an introductory instructional technology course. Paper presented at the *2009 International Conference of the Society for the Information and Technology and Teacher Education*. March 2-6, Charleston, South Carolina.

Schmidt, D., Baran, E., Thompson, A., Mishra, P., Koehler, M.J., & Shin, T. (2009b). Technological Pedagogical Content Knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers. Paper presented at the *2009 Annual Meeting of the American Educational Research Association*. April 13-17, San Diego, California.

Shin, T.S., Koehler, M.J., Mishra, P., Schmidt, D.A., Baran, E., & Thompson, A.D. (2009). Changing Technological Pedagogical Content Knowledge (TPACK) through course evaluations. Paper presented at the *2009 International Conference of the Society for the Information and Technology and Teacher Education*. March 2-6, Charleston, South Carolina.

Shulman, L. S. (1986). "Those who understand: Knowledge growth in teaching." *Educational Researcher Feb. 1986*: 4-14. (AERA Presidential Address).

Swain, C. (2006). Preservice teachers' self-assessment using technology: Determining what is worthwhile and looking for changes in daily teaching and learning practices. *Journal of Technology and Teacher Education, 14*, 29-59.

van Driel, J. H., Verloop, N., & de Vos, W. D. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*(*6*), 673-695.

Vannatta, R.A. & Breyerbach, B. (2000). Facilitating a constructivist vision of technology integration among education faculty and preservice teachers. *Journal of Research on Computing in Education*, *33(2)*, 132-148.

Wang, L., Ertmer, P. A. & Newby, T. J. (2004). Increasing preservice teachers' self-efficacy beliefs for technology integration. Journal of Research on Technology in Education, 36(3), 231-250.

Appendix

Table 19
Self-efficacy Items Related to Technological Knowledge

Item Code	Item Description
TK1	Send an email with an attachment.
TK2	Create a basic presentation using PowerPoint or a similar program.
TK3	Search the internet for information you need.
TK4	Install a new program you would like to use.
TK5	Learn a new program on your own.
TK6	Create a class website, blog, or wiki.

Table 20
Self-efficacy Items Related to Pedagogical Knowledge

Item Code	Item Description
PK1	Motivate students to want to learn in the classroom.
PK2	Manage a classroom with 20-30 students.
PK3	Use knowledge of child development to create age-appropriate lessons.
PK4	Involve students in active learning in the classroom.

Table 21
Self-efficacy Items Related to Content Knowledge

Item Code	Item Description
CK1	How confident are you that you currently know the content in the Utah Core Curriculum Objective listed above? (A math objective is listed above this question)
CK2	How confident are you that you currently know the content in the Utah Core Curriculum Objective listed above? (A language arts objective is listed above this question)
CK3	How confident are you that you currently know the content in the Utah Core Curriculum Objective listed above? (A science or social studies objective is listed above this question)

Table 22
Self-efficacy Items Related to Technological Pedagogical Knowledge

Item Code	Item Description
TPK1	Use digital technologies to motivate students to want to learn in the classroom.
TPK2	Effectively manage a technology-rich classroom.
TPK3	Use knowledge of child development to create age-appropriate lessons that use digital technologies.
TPK4	Use digital technologies to engage students in active learning in the classroom.
TPK5	Use digital technologies to improve my teaching productivity.
TPK6	Use digital technologies to improve the presentation of information to learners.
TPK7	Use digital technologies to help in assessing student learning.

Table 23
Self-efficacy Items Related to Pedagogical Content Knowledge

Item Code	Item Description
PCK1	Teach math using content-specific methods (like inquiry, standards-based math, etc).
PCK2	Teach language arts using content-specific methods (like balanced literacy, etc).
PCK3	Teach social studies using content-specific methods (like using democratic principles, primary source materials, etc).
PCK4	Teach science using content-specific methods (like full inquiry, guided inquiry, etc).

Table 24
Self-efficacy Items Related to Technological Content Knowledge

Item Code	Item Description
TCK1	Use technologies used in mathematics (like calculators, tools for creating graphs and charts, etc).
TCK2	Use technologies used by writers (like tools for desktop publishing, online publishing, etc).
TCK3	Use technologies used in social studies (like digital time lines, geographical information systems, primary source documents, etc).
TCK4	Use technologies used in science (like tools for collecting and analyzing data, etc).

Table 25

Self-efficacy Items Related to Technological Pedagogical and Content Knowledge

Item Code	Item Description
TPACK1	Select appropriate technologies to improve student learning of a topic that is difficult for students to understand.
TPACK2	Use technology in a way that supports content area exploration and learning among your students.
TPACK3	Evaluate computer technology for its fitness for teaching core content in your classroom.
TPACK4	Identify online resources for students to use that support teaching specific topics in your core curriculum.
TPACK5	Use technology to teach math using content-specific methods (like inquiry, standards-based math, etc).
TPACK6	Use technology to teach language arts using content-specific methods (like balanced-literacy, etc).
TPACK7	Use technology to teach social studies using content-specific methods (like democratic principles, primary source materials, etc).
TPACK8	Use technology to teach science using content-specific methods (like full inquiry, guided inquiry, etc).