

Examining the validity of the technological pedagogical content knowledge (TPACK) framework for preservice chemistry teachers

Feng Deng

South China Normal University

Ching Sing Chai

Nanyang Technological University

Hyo-Jeong So

Ewha Womans University

Yangyi Qian

South China Normal University

Lingling Chen

South China Normal University

While various quantitative measures for assessing teachers' technological pedagogical content knowledge (TPACK) have developed rapidly, few studies to date have comprehensively validated the structure of TPACK through various criteria of validity especially for content specific areas. In this paper, we examined how the TPACK survey measure is aligned with the TPACK lesson plan measure and how they are related to the measure of epistemological beliefs about chemistry. The participants were 280 Chinese preservice chemistry teachers enrolled in a university in China. Both exploratory and confirmatory factor analyses were performed on the TPACK survey measure to help to establish validity, including considerations for convergent and discriminant validity. This was followed by the invariance test to examine factorial validity as related to gender. To establish the predictive validity of TPACK, the relationships among teachers' epistemological beliefs, TPACK, and their capacity for planning technology-integrated lessons were also examined. Overall, the results showed that all four types of validity looked at in this study (i.e., convergent, discriminant, factorial, and predictive) were satisfactorily established. Implications for TPACK research and teacher education are also discussed.

Introduction

Since 2005, the framework of technological pedagogical content knowledge (TPACK) has been widely applied for the design and evaluation of teacher professional development courses. (Chai, Koh, & Tsai, 2013; Voogt, Fisser, Roblin, Tondeur, & van Braak, 2013). As some scholars argued, the TPACK framework was powerful not only in unpacking and interpreting knowledge variables that may influence teachers' capacity for integrating information and communication technologies (ICT) within a discipline area, but also in facilitating teachers' implementation of ICT-integrated lessons and relevant assessments (e.g., Angeli & Valanides, 2009; Archambault & Barnett, 2010; Cox & Graham, 2009; Harris, Grandgenett, & Hofer, 2010).

Given the importance of the TPACK framework, a great number of quantitative studies have been devoted to developing various instruments to measure teachers' TPACK. However, much work is still required for enhancing the validity of the TPACK instruments for multiple reasons. First, more research is needed to address some methodological issues detrimental to the validity of the TPACK instruments, such as lack of subject-specificity in TPACK, ambiguous wording of items, unclear definitions and examples of the TPACK components, small number of items per construct (e.g., fewer than three items) and small sample size (e.g., Chai, Koh, & Tsai, 2016; Cox & Graham, 2009).

Second, unsatisfactory convergent validity and discriminant validity were found when the TPACK framework was examined with multiple measures such as survey and rubrics (see Kopcha, Ottenbreit-Lefwich, Jung, & Baser, 2014). Specifically, convergent validity refers to the extent to which indicators of a specific construct converge or share a high proportion of variance in common, whereby discriminant validity refers to the extent to which a construct is truly distinct from other constructs, often using correlation-based statistical strategies (Christensen, 2004). Adequate convergent validity and discriminant validity are important indicators of a satisfactory construct validity, which refers to whether a set of measured variables actually represents the theoretical latent construct those variables are designed to measure (Hair, Black, Babin, & Anderson, 2010).

Third, although several studies (e.g., Chai, Ng, Li, Hong, & Koh, 2013; Lee & Tsai, 2010; Lin, Tsai, Chai, & Lee, 2013) have established the convergent and/or discriminant validity through exploratory/confirmatory factor analyses, little research has further examined the factorial validity, meaning the extent to which a factor structure (e.g., seven component structure) represents a specific construct (e.g., TPACK), and more importantly, whether the construct with such a factor structure has been equivalently measured across different groups (e.g., males and females) of the whole sample (Hair et al., 2010). Fourth, previous studies have seldom examined the TPACK framework through predictive validity, which refers to the extent to which the procedures or behavioral measures can be used to infer or predict certain future behaviors or performance (Christensen, 2004). For example, is a teacher's subject-specific TPACK associated with her belief about the subject itself and her teaching practice (e.g., lesson design)? With this backdrop on the validity issue of TPACK instruments, this study aimed to examine four types of validity (i.e., convergent, discriminant, factorial, and predictive) of the TPACK framework. Only by establishing them will the results yielded from the TPACK instruments be convincing and helpful for TPACK researchers and practitioners.

Literature review

Components of TPACK: A dimensional lens

As an extended form of Shulman's (1986) work on pedagogical content knowledge (PCK), TPACK highlights technological knowledge (TK), in conjunction with pedagogical knowledge (PK) and content knowledge (CK), as another basic form of teachers' professional knowledge (Angeli & Valanides, 2005). Despite the recognised difficulty in defining TK (Koehler & Mishra, 2009), TK generally refers to knowledge of integrating various technologies (Koh, Chai, & Tsai, 2010). PK, on the other hand, can be understood as "teachers' deep knowledge about the processes and practices or methods of teaching and learning" (Koehler & Mishra, 2009, p. 397), and CK refers to knowledge about specific subject matter.

As Mishra and Koehler (2006) proposed, the interactions and intersections of these three knowledge bases (TK, PK, and CK) gave rise to the other four forms of knowledge, namely technological pedagogical knowledge (TPK), technological content knowledge (TCK), pedagogical content knowledge (PCK), and TPACK. Specifically, TPK represents knowledge of using technologies for pedagogical activities (Cox & Graham, 2009), while TCK refers to knowledge of representing subject matter with various technologies (Koh et al., 2010). PCK can be treated as knowledge of how to teach specific subject matter (Shulman, 1986). Finally, TPACK refers to teachers' knowledge of synthesising and transforming the three types of knowledge (TK, PK, and CK) for a particular instructional context (Chai, Koh, & Tsai, 2011).

Generally, the aforementioned seven TPACK components/dimensions (i.e., TK, PK, CK, TPK, TCK, PCK, and TPACK) are more-or-less independent of one another. As implied in the original work of Mishra and Koehler (2006), these dimensions are supposed to be conceptually distinct and mutually exclusively from one another, though certain levels of correlations among them are expected. However, some researchers argued that the boundaries of the TPACK components were unclear, and more clarity about the constructs was requested accordingly (e.g., Archambault & Barnett, 2010; Cox & Graham, 2009). Further discussion about issues pertaining to construct validity is elaborated next.

Validation of TPACK: A multifaceted lens

To assess the aforementioned seven components/dimensions of TPACK among teachers, a group of researchers have developed various instruments with much emphasis on the validity of the TPACK

framework. Regardless of the measures used, researchers seemed to focus heavily on establishing the construct validity, which is situated in convergent and discriminant validity (Hair et al., 2010). In practice, convergent validity can be established through factor loading per item ($>.50$), average variance extracted (AVE) per construct ($>.50$), and construct reliability ($>.70$). Similarly, some researchers tended to establish convergent and discriminant validity through the cross-validation among different measures such as survey, lesson plans, and interviews (e.g., Cavanagh & Koehler, 2013; Hofer & Grandgenett, 2012). Relevant studies were guided by two assumptions (Kopcha et al., 2014; Trochim & Donnelly, 2008): (a) since different measures were built from the same TPACK framework, similar TPACK dimensions should be strongly related (e.g., a high score on TCK on one measure should correspond with a high score on TCK on another measure); and (b) dissimilar dimensions should be mutually exclusive from one another (e.g., a TCK score on one measure should emerge more distinctly from a TPACK score on the same or similar measure).

Among the literature, evidence regarding the construct validity is inconclusive. On one hand, many studies have consistently indicated a lack of satisfactory convergent and discriminant validity. Studies involving multiple measures of TPACK appeared to demonstrate different levels of misalignment among the measures (e.g., Agyei & Keengwe, 2014; Hofer & Grandgenett, 2012; So & Kim, 2009). For example, Kopcha and colleagues (2014) examined TPACK among 27 preservice teachers mainly using the survey of Schmidt et al. (2009) and the content analysis of lesson plans with the rubric developed by Harris et al. (2010). Based on the Spearman rank order correlation on the two measures, they identified low levels of convergence within similar TPACK dimensions (e.g., TPK) and a lack of discrimination between dissimilar dimensions (e.g., TPK and TPACK). Similarly, by adapting a popular survey instrument by Schmidt and colleagues (2009), a group of researchers have reported similar difficulty in differentiating the seven TPACK components (e.g., Archambault & Barnett, 2010; Jang & Tsai, 2012). That is, some components (e.g., TPK, TCK, and TPACK) were unexpectedly merged as new factors/dimensions (see Koh et al., 2010). These researchers argued that the issue of low discriminant validity could be mainly due to the lack of clear definitions and examples of the TPACK components and/or lack of subject-specificity in TPACK (e.g., Chai, Chin, Koh, & Tan, 2013; Chai et al. 2016).

By addressing the above issue, Chai and colleagues have managed to identify the seven TPACK components/dimensions and have reported adequate convergent and discriminant validity (e.g., Chai, Chin et al., 2013; Lin et al., 2013). Specifically, the TPACK measures used in these studies include carefully crafted items to effectively distinguish the seven components. To illustrate, a PCK item was phrased as “Without using technology, I know how to select effective teaching approaches to guide student thinking and learning in my second teaching subject.” (see Chai, Koh, & Tsai, 2011, p. 599). The seven TPACK components were found to be correlated with but distinct from one another (e.g., Chai, Ng et al., 2013; Koh, Chai, & Tsai, 2014). For example, Lin et al (2013) adapted Schmidt et al.’s (2009) survey to investigate subject-specific TPACK among 222 Singaporean science teachers. Confirmatory factor analysis was conducted to *verify* the seven-factor structure and yielded rather high factor loadings ($>.70$), construct reliability ($>.70$), and AVE values (ranging from .67 to .82). Given inconsistent evidence on the construct validity of TPACK, this study adapts relevant instruments by Chai’s research team (e.g., Chai et al., 2011, Chai, Chin et al., 2013) to examine preservice teachers’ subject-specific (chemistry) TPACK.

Although the prior research studies have successfully established the convergent and discriminant validity through factor analysis, further rigorous tests (e.g., test of measurement invariance) are still required to enhance the factorial validity of TPACK and its representation in assessments. In practice, factorial validity can be indicated through satisfactory model-fit indexes of a model with specific factor structure and the accomplishment of measurement invariance. The former was evident in several existing studies (e.g., Koh et al., 2014; Lin et al., 2013), while the latter has been hardly recognised. That is, few researchers have explicitly examined whether the seven-factor structure of TPACK was equivalent for different groups (e.g., gender, ethnics, and teaching experience) involved in relevant participant populations. This constitutes a research gap that the current study aimed to fill.

Furthermore, as noted above, some researchers (e.g., Agyei & Keengwe, 2014; Hofer & Grandgenett, 2012; Kopcha et al., 2014) emphasised the validation of TPACK across multiple measures. In some sense, their work at least implicitly suggested the importance of another type of validity, predictive validity, meaning the extent to which the measures concerned can be used to infer or predict certain related behaviours or performance (Christensen, 2004). In the case of TPACK, this would mean that the TPACK components

should be able to predict relevant instructional behaviours such as planning an ICT-integrated lesson. In practice, this can be established via correlation and/or regression between assessments. Given a dearth of research studies have explicitly examined the predictive validity, this study attempts to address such a gap in the TPACK research.

Development of TPACK: A presage-process-product lens

Apart from ensuring the multifaceted validity of TPACK, researchers can move a step forward by using a presage-process-product (3P) lens (Biggs, 1993) to interpret the complex nature of TPACK. As its name implies, the 3P model consists of three parts: presage, process, and product. In the context of teacher knowledge research, presage can be understood as teachers' perceptions of their work contexts and their beliefs about teaching, learning, and the nature of subject matter knowledge. Process can refer to TPACK, which has been characterised as a dynamic form of knowledge that is associated with teachers' beliefs and practice (Angeli & Valanides, 2009; Chai, Chin et al., 2013). The TPACK measure underlies the design process. Because it is not quite possible to derive an objective measure of the complex design processes that each individual preservice teacher was engaged in, we used the TPACK measure as a proxy to the design process (Chai, Tan, Deng, & Koh, 2017). The process feature of TPACK also gains support from the epistemological resources framework by Hammer and Elby (2002). That is, TPACK can be treated as teachers' enactment of TK, PK, CK and other overlapping constructs as epistemological resources to formulate ICT-integrated lessons for a specific group of students (Chai, Koh et al., 2013). Product refers to teachers' planning, implementing, and reflecting on their lesson activities in practice, which can be understood as "TPACK-in-action" (Koh, Chai, & Tay, 2014, p. 20). For example, teachers' lesson plans can be considered as a product because they can demonstrate the results of teachers' decision-making and pedagogical reasoning (Harris et al., 2010). Inspired by the 3P model, this study explored the relationships among preservice teachers' epistemological beliefs about chemistry (presage), TPACK (process), and their lesson planning (product). By doing so, we hoped: (a) to provide empirical evidence for the 3P model, (b) to interpret TPACK from a more comprehensive and systemic perspective, and (c) to enhance the validity of TPACK through associated measures.

Research purposes and questions

Based on the literature review, the current study aimed to examine the convergent, discriminant, factorial, and predictive validity with two established and validated measures of TPACK: Chai et al. (2011) survey and Harris et al. (2010) rubric. Specifically, this study was guided by three research questions:

1. Can the convergent and discriminant validity of the TPACK survey be established with the Chinese preservice chemistry teachers?
2. If so, can the factorial validity of the TPACK survey be established?
3. If so, can the predictive validity be established with the epistemological belief measure, the TPACK survey and the lesson plan rubric?

The three research questions are elaborated as follows. The first question focuses on whether the seven-factor structure of the TPACK survey measure can be recognised through exploratory and confirmatory factor analyses. The second question seeks to understand whether the verified seven-factor structure is equivalently measured across gender. Regarding the characteristics of the participants in this study, they shared many similarities in age (i.e., 21-22 years old), ethnics (i.e. Chinese) and teaching experience (i.e., preservice). In this vein, gender seemed to be the main reference for dividing the participants into different groups (i.e., male and female) for the invariance test to examine the factorial validity. Unlike other studies (e.g., Chai et al., 2016; Koh et al., 2014) that investigated the gender difference in the mean scores on the seven TPACK components, the current study instead examined the invariance of the factor structure across gender. This was due to the main focus (e.g., validity issue) of this study and the coherence among the three research questions. With the factorial validity established and through the 3P lens, the third research question investigates the relationships among participants' epistemological beliefs about chemistry, TPACK, and lesson planning.

Methods

Participants

The participants of this study were 280 preservice chemistry teachers (54% female, 46% male) who just started their sixth semester of a 4-year program during Feb 2016. As the program required, these preservice chemistry teachers were expected to complete various chemistry courses including inorganic chemistry, organic chemistry, analytic chemistry, and physical chemistry during the first 3 years. Generally, three ICT-related courses were provided for the participants. Specifically, one compulsory course was available during the first semester and it covered general aspects of technologies such as Word, PowerPoint, Excel, and Access. The other two optional courses were scheduled at the third and fifth semester, respectively. They both involved the use of particular software programs (e.g., video editing and producing, Flash, and online chemistry game) for chemistry teaching (e.g., using Flash to represent the process of certain chemical reaction at a submicroscopic level). During the sixth semester, these student teachers had the opportunity to take another optional course (i.e., Technological Pedagogical Content Knowledge for Chemistry) which highlighted the TPACK framework and its application in chemistry teaching. Before the course, all the participants had completed three teacher education courses on curriculum, teaching, and learning chemistry, as well as chemistry teaching skills. They had gained relevant experiences in designing lesson activities and micro-teaching in the previous courses. At the time of the study, the participants had no classroom teaching experience since the teaching practicum was scheduled at the last year of the preservice education program.

Instrumentation

Three instruments were used in this study: an adapted Chai et al.'s (2011) TPACK survey, Harris et al.'s (2010) rubric for TPACK integration, and an adapted instrument (Deng, Chai, Tsai, & Lin, 2014) for assessing epistemological beliefs about chemistry. Compared to the original instrument, the adapted TPACK questionnaire: (a) specified relevant content-related components (CK, TCK, PCK, and TPACK), (b) included only one CK component rather than two (first- and second-teaching subjects), (c) used technological terms more familiar or chemistry-relevant to the participants, and (d) refined items with relatively low factor loading. As a result, the adapted survey consisted of a total of 24 items, with 3 or 4 items for each TPACK component. Sample items are presented in Table 1. All the items were measured on a 7-point Likert-type scale (1 = *strongly disagree*, 7 = *strongly agree*). Based on the data in this study, the Cronbach's α values ranged from .79 to .93, suggesting acceptable internal consistency (Hair et al., 2010).

Table 1
Sample items of the seven TPACK components

Components	Sample Items
TK (4 items)	I can learn technology easily.
PK (3 items)	I am able to help my students to monitor their own learning.
CK (3 items)	I have sufficient knowledge about chemistry.
TPK (3 items)	I am able to facilitate my students to use technology to plan and monitor their own learning.
TCK (4 items)	I know about the technologies that I have to use for the research of chemistry.
PCK (3 items)	Without using technology, I can help my students to understand chemistry through various ways.
TPACK (4 items)	I can teach lessons that appropriately combine chemistry, teaching approaches and technologies (e.g., Flash, Virtual Chemistry Lab).

Harris et al. (2010) developed and validated the TPACK-based technology integration assessment rubric to assess TPACK inferred from teaching artifacts such as lesson plans created by student teachers. This rubric can be employed to rate lesson plans on four dimensions: (a) alignment between curriculum goals and technologies (TCK), (b) alignment between instructional strategies and technologies (TPK), (c) the appropriateness of technology selected within the planning product (TPACK), and (d) *fit* of content, pedagogy, and technology. Each dimension was scored on a 4-point scale, with "4" standing for the highest level of alignment, appropriateness (e.g., exemplary), and fit. In this study, two trained raters used the rubric

to code 56 (20%) of the lesson plans randomly selected from the participants. Cohen's kappa was computed to check the inter-rater reliability, and the resulting kappa was .87.

Two subscales from an instrument (Deng et al., 2014) on scientific epistemological beliefs were also adapted for assessing the predictive validity in this study. The two subscales were (a) "beliefs about the constructive-inventive source of chemistry knowledge" (3 items) and (b) "beliefs about the tentative nature of chemistry knowledge" (4 items). Sample items for these two subscales included "Chemistry knowledge comes from chemists' imagination and creativity" and "Chemistry knowledge is not unchangeable", respectively. The Cronbach's α values for the two subscales were .92 and .90, individually, indicating very good internal consistency (Hair et al., 2010).

At the first session of the course, the instructor (the first author) briefly introduced the goals of the course and expectations for the preservice chemistry teachers. He then administered the survey instrument that consisted of the seven TPACK subscales and the two subscales of epistemological beliefs. Before the participants started completing the questionnaire, the instructor guided them to go through each item and provided relevant explanation about its meaning when necessary. On average, the participants completed the questionnaire within 25 minutes. They were asked to submit an ICT-integrated lesson plan by the next course session from which TPACK was officially introduced and elaborated. The lesson plan typically consisted of instructional goals, teaching strategies, activities of the teacher and students, and relevant rationale for the design. Participation was voluntary for both the survey and lesson planning.

Data analysis

To answer the first research question, principal component analysis (PCA) with the Varimax rotation method was conducted to investigate the seven-factor structure of the instrument. This was followed by a confirmatory factor analysis (CFA) to further verify the seven-factor structure. As Hair et al. (2010) suggested, a combination of multiple fit indexes should be used to examine how the model fits the data. Specifically, five indexes were used in this study: χ^2/df (< 3.0), CFI ($> .90$), TLI ($> .90$), RMSEA ($< .07$), and SRMR ($< .08$). Factor loadings generated from both PCA and CFA can be used to examine the convergent validity, while the model-fit test results would provide valuable information on the discriminant validity of TPACK. Next, standardised factor loadings were then used to calculate the values of AVE and construct reliability, which are the other two indicators of convergent validity. Specifically, AVE was calculated as the total of all squared standardised factor loadings divided by the number of items; while construct reliability was computed from the squared sum of factor loadings for each construct and the sum of the error variance terms for a construct (see Hair et al., 2010). Finally, the discriminant validity was examined by testing whether the square roots of AVEs were greater than the correlation estimates.

To answer the second research question, multi-group analyses with maximum likelihood method were then employed to test the measurement invariance across gender. According to Hair et al. (2010), the measurement invariance should be tested at each of the following steps: (a) configural invariance (i.e., invariance of factor structure), (b) metric invariance (i.e., invariance of factor loadings), (c) scalar invariance (invariance of intercepts and means), (d) factor covariance invariance (i.e., invariance of covariance between constructs), (e) factor variance invariance (invariance of variances of the constructs), and (f) error variance invariance (i.e., invariance of error variance in the indicators). Generally, each *neighbouring* pair of models in the above sequence (i.e., from (a) to (f)) is nested because different sets of parameters are constrained to be equal across groups (e.g., gender and grade levels) in the more restricted model (Chen, Sousa, & West, 2005). Following Chen et al.'s suggestion, four criteria were used to compare the fit for two nested models in this study: (a) the chi-square difference ($\Delta \chi^2$) test, (b) Δ CFI test, (c) Δ TLI test, and (d) Δ RMSEA test. Considering that the use of $\Delta \chi^2$ test has been criticised due to its sensitivity to sample size (Hair et al., 2010), the other three criteria were thus used for the triangulation purpose. Specifically, Δ CFI $\leq .01$, Δ TLI $\leq .022$, and Δ RMSEA $\leq .015$ suggested non-significant changes in model fit for testing measurement invariance, respectively.

To answer the third research question, similar to Kopcha et al.'s (2014) study, Spearman rank order correlation was conducted to examine the associations among scores on three TPACK components (i.e., TCK, TPK, and TPACK) and scores on the four-dimension rubric by Harris et al. (2010). Besides, path analysis was performed on the data of the 56 randomly selected participants to investigate any significant

relations among preservice teachers' epistemological beliefs about chemistry, TPACK (as a whole), and their lesson planning (represented through the total scores on the four dimensions of the rubric).

Results

Convergent and discriminant validity of TPACK

As shown in Table 2, the results of PCA showed that seven factors were identified and explained 78.36% of total variance, with each explaining 9.12% to 14.54%. The factor loadings of all items were greater than .50. Similarly, all the standardised factor loadings produced from CFA were larger than .50. These jointly indicated acceptable convergent validity at the item level.

Table 2
Means, standard deviations, factor loadings (λ), AVEs, and construct reliability

TPACK components (percentage of variance explained)	<i>M</i>	<i>SD</i>	λ -PCA	λ -CFA	AVE ^a	Construct reliability ^b
Factor 1: TPACK (14.54%)	4.83	1.18			.80	.94
TPACK1	4.76	1.32	.84	.90		
TPACK2	4.89	1.18	.84	.89		
TPACK3	4.80	1.27	.84	.90		
TPACK4	4.87	1.24	.83	.90		
Factor 2: TCK (12.45%)	4.48	1.15			.64	.88
TCK1	4.52	1.27	.80	.80		
TCK2	4.34	1.36	.79	.89		
TCK3	4.59	1.26	.73	.74		
TCK4	4.48	1.50	.70	.77		
Factor 3: TK (12.14%)	4.62	1.05			.62	.87
TK1	4.63	1.26	.78	.82		
TK2	4.31	1.32	.76	.69		
TK3	4.64	1.14	.75	.87		
TK4	4.90	1.16	.71	.77		
Factor 4: TPK (10.22%)	4.49	1.14			.71	.88
TPK1	4.55	1.26	.84	.85		
TPK2	4.50	1.30	.78	.83		
TPK3	4.42	1.25	.78	.84		
Factor 5: PCK (10.18%)	4.65	.94			.68	.87
PCK1	4.65	1.02	.86	.89		
PCK2	4.61	1.14	.80	.80		
PCK3	4.71	1.03	.72	.80		
Factor 6: CK (9.70%)	4.40	1.17			.67	.86
CK1	4.29	1.37	.79	.83		
CK2	4.60	1.24	.79	.81		
CK3	4.30	1.35	.75	.82		
Factor 7: PK (9.12%)	4.50	.95			.57	.86
PK1	4.31	1.22	.79	.80		
PK2	4.52	1.05	.73	.73		
PK3	4.67	1.11	.73	.74		

Notes. The factor loadings of each item on the other factors were generally less than .30, and they were not presented here for simplicity.

a $AVE = \sum \lambda^2 / n$, n is the number of items per construct; b Construct Reliability = $(\sum \lambda^2) / ((\sum \lambda^2) + (\sum (1 - \lambda^2)))$.

The CFA results also showed that the seven-factor structure expressed satisfactory model-fit for the whole sample ($\chi^2 = 392.26$, $df = 231$, $\chi^2/df = 1.70$, IFI = .966, CFI = .966, TLI = .959, RMSEA = .050, SRMR = .039). This suggested that the seven components were appropriate for assessing Chinese preservice chemistry teachers' TPACK, and they were correlated but distinct from one another. Hence, acceptable discriminant validity could be initially established. Furthermore, all the values of AVEs and construct

reliability were larger than .50 (Table 2), suggesting adequate convergence at the construct level (Hair et al., 2010). As seen in Table 3, all of the square roots of AVEs were greater than the off-diagonal elements in the corresponding rows and columns, indicating adequate discriminant validity of the instrument (see Deng et al., 2014).

Table 3
Correlations among and discriminant validity of the seven TPACK components

Components	TK	PK	CK	TPK	TCK	PCK	TPACK
TK	(.79)						
PK	.51	(.75)					
CK	.46	.46	(.82)				
TPK	.54	.51	.41	(.84)			
TCK	.49	.45	.55	.40	(.80)		
PCK	.48	.46	.51	.36	.48	(.82)	
TPACK	.49	.43	.54	.50	.59	.49	(.89)

Note. All correlations (off-diagonal figures) are significant at the .01 level; Diagonal figures in the parentheses are the square root of AVE from items.

Table 4
Fit statistics for testing measurement invariance across gender

Model#	χ^2	df	RMSEA	CFI	TLI	Model comparison	$\Delta \chi^2$	Δdf
Model 1	756.32	462	.048	.939	.927	--	--	--
Model 2	778.43	479	.047	.938	.929	2 v 1	22.11	17
Model 3	799.15	503	.046	.939	.933	3 v 2	20.72	24
Model 4	840.56	524	.047	.935	.931	4 v 3	41.41	21
Model 5	855.95	531	.047	.933	.930	5 v 4	15.39	7
Model 6	890.03	555	.047	.931	.931	6 v 5	34.08	24

Note. N = 280, 151 (females) v 129; Model 1- configural invariance; Model 2 - metric invariance; Model 3 - scalar invariance; Model 4 – factor covariance invariance; Model 5 –factor variance invariance; Model 6 – error variance invariance. * p <.05

Factorial validity of TPACK

According to the test of the measurement invariance, the seven-factor structure was equivalently measured across gender. As Table 4 shows, the chi-square difference was not significant ($p > .05$) between any two neighboring nested models. Using Hair et al.'s (2010) criterion, this supported the hypothesis that the seven-factor structure was invariant across gender. Such interpretation was also congruent with the results of other tests; that is, there were no noticeable differences in CFI ($< .01$), TLI ($< .022$) and RMSEA ($< .015$) for the above pairs of models. Therefore, it was concluded that the seven-factor structure was invariant across male and female preservice teachers.

Predictive validity of TPACK

According to the results of Spearman rank order correlation, significant correlations were observed for similar TPACK components. Specifically, the TCK scores on the survey (Chai et al., 2011) was significantly correlated with those on the rubric (Harris et al., 2010), $r = .66, p < .001$. Similarly, significant moderate correlation was also spotted between the TPK scores on the two measures, $r = .62, p < .001$. For the TPACK component, moderate to high correlations were recognised between its scores on the survey and the last two dimensions of Harris et al.'s rubric, namely the *appropriateness* ($r = .70, p < .001$) and *fit* ($r = .64, p < .001$) dimensions. Besides, the mean scores for the four dimensions of the rubric were 2.89 (TCK), 2.98 (TPK), 2.77 (TPACK-appropriateness), and 2.77 (TPACK-fit), respectively. They were generally aligned with the mean scores of relevant TPACK components (see Table 2) such as TCK (4.48), TPK (4.49), and TPACK (4.83). These results not only indicated acceptable predictive validity of TPACK, but also provided evidence of convergence across similar constructs (i.e., convergent validity). These findings, however, were inconsistent with those reported in Kopcha et al.'s study (2014).

Lastly, the results of path analysis showed that the preservice teachers' epistemological beliefs about chemistry significantly predicted their TPACK as a whole, which then predicted their total scores on Harris et al.'s (2010) rubric. In particular, preservice chemistry teachers' beliefs about the constructive-inventive source ($\beta = .43$, $t = 3.42$, $p = .001$) and the tentative nature of chemistry knowledge ($\beta = .43$, $t = 3.41$, $p = .001$) were two significant contributors to their TPACK, $F(2, 53) = 47.07$, $p < .001$, $R^2 = .64$, Adjusted $R^2 = .63$. Their TPACK also significantly contributed to their lesson planning, $F(1, 54) = 75.83$, $p < .001$, $R^2 = .58$, Adjusted $R^2 = .58$. However, scores of the two beliefs subscales did not have significant direct effects on the total scores of the rubric.

Discussion

This study aimed to examine the construct, factorial, and predictive validity of the TPACK framework. Three main findings were obtained in this study. First, acceptable convergent validity and discriminant validity were successfully established. Second, adequate factorial validity was supported by the measurement invariance test across gender. Third, satisfactory predictive validity of TPACK was evident through the significant moderate correlations between the scores of similar TPACK components on the two measures: Chai et al. (2011) survey and Harris et al. (2010) rubric. The following discussions highlight the above findings and unfold their implication for theory and practice for teacher education.

Reinforcing multifaceted validity of TPACK

This study supported the findings of several empirical studies (e.g., Chai et al., 2011; Chai, Chin et al., 2013; Koh et al., 2014; Lin et al., 2013) in terms of construct validity. Congruent with the framework by Mishra and Koehler (2006), the seven TPACK components were found to be correlated but distinct from one another. This may also help address others' (e.g., Archambault & Barnett, 2010; Cox & Graham, 2009) concern on the unclear boundaries of the TPACK components. However, such finding was not aligned with other studies (e.g., Hofer & Grandgenett, 2012; Kopcha et al., 2014) that reported unsatisfactory convergent and discriminant validity. For example, despite using instruments and analytic approaches similar to Kopcha et al.'s study, opposite findings were noticed. Such apparent inconsistency could be explained in at least three ways. First, lack of subject-specificity in Schmidt et al. (2009) survey might reduce the construct validity (Chai et al., 2016). As Kopcha and colleagues argued, the "Schmidt et al. (2009) survey items internally lacked specificity in an effort to improve application across contexts" (p.94). However, the adapted TPACK instrument (Chai et al., 2011) used in this study questionnaire underscored the subject-specific nature of TPACK and all the participants were majored in the same subject (chemistry), which might help enhance the correlation between scores on the two measures (i.e., survey and rubric). Second, inadequate clarity or unclear definitions of the TPACK components may be another possible account. Compared to other studies (e.g., Archambault & Barnett, 2010; Jang & Tsai, 2012), the current study might put more effort in crafting the items to effectively distinguish the TPACK components, especially PCK and TPACK. Third, cultural difference could be considered for interpreting the inconsistent findings. As discussed by Kopcha et al., the participants in their study might have overestimated themselves when reporting on their own TPACK profiles (see also Kopcha & Sullivan, 2007). In comparison, perhaps influenced by the Confucian culture that esteems humility (Shafer, Fukukawa, & Lee, 2007), the preservice teachers in mainland China appeared to be humbler and they might even under-represent themselves when reporting on their knowledge efficacy. All the above three explanations are hypothetical only, and require further warrant by research. For example, researchers may conduct similar studies within different subject-specific contexts (e.g., physics, biology, and mathematics), and examine the construct validity through the two common measures (survey and rubric) within different cultural backgrounds (e.g., East and West). Given the culture-specific nature of TPACK (Chai et al., 2016), future studies can further explore how various cultural variables may shape teachers' TPACK and its development.

The current study also demonstrated how to help to establish the factorial validity of TPACK and its representation in assessments through the test measurement invariance. Within the research field of TPACK, tests of measurement invariance appear essential, especially when the participant population involves at least two subgroups (e.g., gender, subject, and cultural background). For example, for studies that aim to investigate differences in TPACK between two groups (e.g., males and females), measurement invariance must be tested to check whether specific factor structures can be equivalently measured for different groups. As Hair et al. (2010) suggested, the invariance needs to be established at least at the level of scalar invariance (see Model 3 in Table 4), which tests for the equality of the measured variable intercepts (i.e.,

means) on the construct. Without the guarantee of scalar invariance, any comparison across groups would be problematic. In this study, the data supported for the measurement invariance even at the level of error variance invariance (see Model 6 in Table 4). This suggested the appropriateness of using the same instrument (with seven constructs and 24 items) for measuring the TPACK of male and female teachers. Future studies may need to consider the tests of measurement invariance when participants involve different subgroups such as preservice/novice versus in-service/experienced teachers, primary vs. secondary teachers, as well as teachers teaching different subjects.

Leveraging multiple measures for TPACK

Two common measures (survey and rubric) were used in this study, and the results of Spearman rank order correlation suggested that these two measures were mutually supported. This finding is consistent with several studies that reported different levels of misalignment among various measures (e.g., Agyei & Keengwe, 2014; Hofer & Grandgenett, 2012). Similar to the Kopcha et al.'s (2014) study, these studies employed Schmidt et al. (2009) survey as one of the important measures. Apart from the aforementioned lack of specificity, Schmidt et al.'s survey may be confronted with another methodological issue that factor analyses were performed on each TPACK component rather than all the items. This may help explain the phenomenon that the low level of associations was recognised between Schmidt et al.'s (2009) survey and other measures (e.g., interview), whereas some other measures could triangulate one another (see Hofer & Grandgenett, 2012).

The plausible mismatch across different measures should not hinder researchers' confidence in the TPACK assessment, since a certain level of disagreement between any two measures is expected. Instead, future studies may consider the use of multiple measures for assessing TPACK, which may allow for better understanding and even necessary refinement/revision of the TPACK framework. As evident in the literature, researchers have developed various instruments for examining teachers' TPACK including professed and enacted TPACK. The professed TPACK, often decontextualised, was usually assessed through surveys (e.g., Likert-type scales), open-ended questionnaires, and interviews (Chai, Koh et al., 2013). To assess teachers' enacted TPACK, often more context-sensitive, researchers tend to rely on other performance-oriented assessments such as class observation, lesson plans, design tasks, learning activities, and case development (see Archambault, 2016 for a review). As Chai et al. (2016) suggested, developing appropriate tests and rubrics for evaluating teachers' TPACK-in-action (Koh et al., 2014) other than their self-report of TPACK deserves more attention from researchers. Similarly, some researchers (e.g., Abbott, 2011; Koehler, Shin, & Mishra, 2012; Kopcha et al., 2014) also proposed that future research should place more emphasis on measuring the processes of TPACK construction (enacted TPACK), which may require other measures such as content and discourse analyses of teachers' design talk, authentic tasks, reflection, portfolios, and other teaching artifacts.

Unpacking theoretical connection to TPACK

As found in this study, preservice chemistry teachers' TPACK was significantly correlated with their epistemological beliefs about chemistry and lesson planning. This finding generally lent support to Biggs's (1993) 3P model that underlines the relations among presage, process, and product. In this study, teachers' beliefs (presage) about the constructive-inventive source and the tentative nature of chemistry knowledge both significantly predicted their TPACK (process) as a whole. This could be interpreted through the nature of TPACK that it can be seen as a form of *dynamic* knowledge creation (Chai et al., 2016). Besides, from the perspective of epistemological resources (Hammer & Elby, 2002), student teachers' epistemological beliefs and TPACK can be considered as an enactment of their integrative and coherent use of relevant epistemological resources (e.g., knowledge is constructed/created stuff, and knowledge is subject to change). As expected, teachers' TPACK was found to significantly contribute to their lesson planning (product) in this study. However, such a relationship should not be taken for granted. This is because possessing adequate TPACK, and even being able to articulate the rationale for making instructional decisions may or may not lead to a teacher's design capacity for planning high-quality ICT-integrated lessons. In other words, there might be a gap between teachers' TPACK and their competency to translate TPACK into lesson designing in authentic contexts. This again highlights the difference and relation between the professed and enacted TPACK. To develop a more fine-grained understanding about their relationship, more empirical studies (especially qualitative studies) are highly required.

Furthermore, it was found that teachers' TPACK may serve as a pivotal role in mediating the relationship between beliefs and lesson planning (product). Teachers' epistemological beliefs about chemistry did not have a significant direct effect on their lesson planning, but was mediated through TPACK. Literature on the *belief-practice* relationship may help interpret this unexpected finding. As many scholars maintained, there is a complex relationship between teachers' beliefs and practice, which may be facilitated or impeded by various external and internal factors (e.g., Basturkmen, 2012; Mansour, 2009). As one of the internal factors, teachers' TPACK may exert much more power on lesson planning when compared with their beliefs. In this sense, the role of beliefs was overtaken by knowledge (Buehl & Beck, 2015). Another possible explanation could be the selectivity of the belief-practice relationship. As some researchers argued, such a relationship may vary based on the types of beliefs under consideration (Pajares, 1992) and their status within teachers' belief system (Green, 1971). In this study, epistemological beliefs may be implicit to the participants since they have seldom been exposed to the learning experience that made epistemology explicit, and they would hardly reflect on their beliefs about the source and nature of (chemistry) knowledge, not to mention to consciously translate their beliefs into practice (e.g., lesson planning). These tentative interpretations would need to be tested in future studies.

Improving teacher education on TPACK

The preservice chemistry teachers reported only fairly adequate TPACK. Specifically, their survey scores on the seven TPACK components ranged from 4.40 (CK) to 4.83 (TPACK) out of 7. These findings could be explained by the fact that this study was conducted at the beginning of the TPACK course. Besides, preservice teachers' rubric scores ranged from 2.77 (TPACK) to 2.98 (TPK) out of 4. This suggested that the participants in some degree were able to design ICT-integrated lessons, but their lesson design capacity and pedagogical reasoning can be further improved. Prior to this course, the participants had completed relevant courses on educational technology, chemistry, and pedagogy. Although they gained certain PCK during the fifth semester, they still lacked experience in how to effectively integrate technology, pedagogy, and chemistry within any courses provided by the university. These may account for the teachers' mean scores on both the two measures. To address the above shortcomings, the following suggestions can be considered.

First, the TPACK framework can be introduced to preservice teachers at the beginning of the teacher education program. This may help preservice teachers to develop a holistic understanding about the TPACK components, which in turn enables them to appreciate the importance of various courses (e.g., technology-, subject-, and pedagogy-oriented) in shaping relevant knowledge (e.g., TK, CK, and PK). Second, teacher educators should be trained to design their course based on how preservice teachers can use the knowledge gained from relevant courses (Tondeur et al., 2012), which in turn allows them to model the TPACK reasoning process (Voogt et al., 2013) to the preservice teachers. That is, preservice teachers need to be taught how to synthesise the basic and derived knowledge components (e.g., TK, PK, CK, TPK, TCK, and PCK) as epistemological resources to achieve satisfactory TPACK. Third, Harris et al.'s (2010) rubric and the TPACK surveys can be utilised to scaffold the development of ICT course for enhancing preservice teachers' TPACK. When developing an ICT course for preservice teachers, more emphasis should be given to Levels 3 and 4 of the four dimensions of the rubric, and/or the TPK, TCK, and TPACK components. For example, the instructor can introduce the preservice teachers to the rubric and/or the relevant items from the TPACK survey, and then encourage them to use the rubric or survey items as a design goal to prepare and polish their lesson plans.

Limitations and future research

This study has some limitations. First, the findings were based on statistical analysis of survey data and content analysis of the lesson plans, while more detailed information was not collected in this study. For example, one-to-one interviews were not conducted, and fine-grained analysis of the alignment or misalignment of each participant's survey and rubric scores were not performed. Future studies may address these issues to probe more deeply into preservice teachers' rationale for their lesson plans, and tease out possible accounts for the alignment or misalignment across the two measures. Second, only one lesson plan was collected from each participant, and perhaps that single lesson plan may not accurately embody a teacher's TPACK. Given the topic-specific nature of PCK/TPACK (Chai et al., 2016; Shulman, 1986), teachers' lesson design capacity may vary across different topics/modules (e.g., chemistry concepts, principles, and elements) of secondary school chemistry. Therefore, future research may consider

replicating this study by collecting multiple lesson plans for different topics of chemistry or even other subjects (e.g., physics, mathematics, and language). Third, the sample size associated with the study was relatively small, which might limit the generalisation of the findings. Future studies may involve a larger population to use different subsamples for PCA and CFA separately, which would achieve the cross-validation purpose (see Deng et al., 2014).

Conclusion

The current study may contribute to demonstrating how to establish the construct, factorial, and predictive validity of the TPACK framework and assessments that strive to represent it. It also provided some empirical evidence that supports the convergent and discriminant validity in this context. Future research should further enhance TPACK assessment in teacher education and fully utilise assessment to guide the development and evaluation of ICT courses. Moreover, researchers may probe deeper to study the specificity of TPACK and its connection with fields of beliefs and practice.

Acknowledgments

The research is supported by South China Normal University (including the Project 2016RCYJKYQD-002 and the Junior Faculty Cultivation Fund). The authors would also like to thank the reviewers and the guest editors for their insightful suggestions.

References

- Abbitt, J. T. (2011). Measuring technological pedagogical content knowledge in preservice teacher education: A review of current methods and instruments. *Journal of Research on Technology in Education*, 43(4), 281-300. <https://doi.org/10.1080/15391523.2011.10782573>
- Agyei, D. D., & Keengwe, J. (2014). Using technology pedagogical content knowledge development to enhance learning outcomes. *Education and Information Technologies*, 19(1), 155-171. <https://doi.org/10.1007/s10639-012-9204-1>
- Angeli, C., & Valanides, N. (2005). Preservice elementary teachers as information and communication technology designers: An instructional systems design model based on an expanded view of pedagogical content knowledge. *Journal of Computer Assisted Learning*, 21(4), 292-302. <https://doi.org/10.1111/j.1365-2729.2005.00135.x>
- 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 & Education*, 52(1), 154-168. <https://doi.org/10.1016/j.compedu.2008.07.006>
- Archambault, L. (2016). Exploring the use of qualitative methods to examine TPACK. In M. C. Herring, M. J. Koehler, & P. Mishra (Eds.), *Handbook of technological pedagogical content knowledge (TPACK) for educators* (pp 65-85). New York, NY: Routledge.
- Archambault, L. M., & Barnett, J. H. (2010). Revisiting technological pedagogical content knowledge: Exploring the TPACK framework. *Computers & Education*, 55(4), 1656-1662. <https://doi.org/10.1016/j.compedu.2010.07.009>
- Basturkmen, H. (2012). Review of research into the correspondence between language teachers' stated beliefs and practices. *System*, 40(2), 282-295. <https://doi.org/10.1016/j.system.2012.05.001>
- Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63(1), 3-19. <https://doi.org/10.1111/j.2044-8279.1993.tb01038.x>
- Buehl, M. M., & Beck, J. S. (2015). The relationship between teachers' beliefs and teachers' practices. In H. Fives, & M. G. Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 66-84). New York, NY: Routledge.
- Cavanagh, R. F., & Koehler, M. J. (2013). A turn toward specifying validity criteria in the measurement of technological pedagogical content knowledge (TPACK). *Journal of Research on Technology in Education*, 46(2), 129-148. <https://doi.org/10.1080/15391523.2013.10782616>
- Chai, C. S., Chin, C. K., Koh, J. H. L., & Tan, C. L. (2013). Exploring Singaporean Chinese language teachers' technological pedagogical content knowledge and its relationship to the teachers'

- pedagogical beliefs. *The Asia-Pacific Education Researcher*, 22(4), 657-666.
<https://doi.org/10.1007/s40299-013-0071-3>
- Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2011). Exploring the factor structure of the constructs of technological, pedagogical, content knowledge (TPACK). *The Asia-Pacific Education Researcher*, 20(3), 595-603.
- Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2013). A review of technological pedagogical content knowledge. *Journal of Educational Technology & Society*, 16(2), 31-51.
- Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2016). A review of the quantitative measures of technological pedagogical content knowledge (TPACK). In M. C. Herring, M. J. Koehler, & P. Mishra (Eds.). *Handbook of technological pedagogical content knowledge (TPACK) for educators* (pp 131-160). New York, NY: Routledge.
- Chai, C. S., Ng, E. M. W., Lee, W. H., Hong, H.-Y., & Koh, J. H. L. (2013). Validating and modeling technological pedagogical content knowledge (TPCK) framework among Asian pre-service teachers. *Australasian Journal of Educational Technology*, 29(1), 41-53. <https://doi.org/10.14742/ajet.174>
- Chai, C. S., Tan, L., Deng, F., & Koh, J. H. L. (2017). Examining pre-service teachers' design capacities for web-based 21st century new culture of learning. *Australasian Journal of Educational Technology*, 33(2), 129-142. <https://doi.org/10.14742/ajet.3013>
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12(3), 471-492.
https://doi.org/10.1207/s15328007sem1203_7
- Christensen, L. B. (2004). *Experimental methodology*. Boston, MA: Pearson Education.
- Cox, S., & Graham, C. R. (2009). Diagramming TPACK in practice: Using an elaborated model of the TPACK framework to analyze and depict teacher knowledge. *TechTrends*, 53(5), 60-69.
<https://doi.org/10.1007/s11528-009-0327-1>
- Deng, F., Chai, C. S., Tsai, C. C., & Lin, T. J. (2014). Assessing South China (Guangzhou) high school students' views on nature of science: A validation study. *Science & Education*, 23(4), 843-863.
<https://doi.org/10.1007/s11191-013-9674-6>
- Green, T. (1971). *The activities of teaching*. New York, NY: McGraw-Hill.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). New York, NY: Prentice Hall.
- Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169-190). New Jersey, NJ: Lawrence Erlbaum Associates.
- Harris, J., Grandgenett, N., & Hofer, M. (2010). Testing a TPACK-based technology integration assessment rubric. In C. D. Maddux, D. Gibson, & B. Dodge (Eds.), *Research highlights in technology and teacher education 2010* (pp. 323-331). Chesapeake, VA: Society for Information Technology & Teacher Education (SITE).
- Hofer, M., & Grandgenett, N. (2012). TPACK development in teacher education: A longitudinal study of preservice teachers in a secondary MA Ed. program. *Journal of Research on Technology in Education*, 45(1), 83-106. <https://doi.org/10.1080/15391523.2012.10782598>
- Jang, S. J., & Tsai, M. F. (2012). Exploring the TPACK of Taiwanese elementary mathematics and science teachers with respect to use of interactive whiteboards. *Computers & Education*, 59(2), 327-338. <https://doi.org/10.1016/j.compedu.2012.02.003>
- Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge. *Contemporary Issues in Technology and Teacher Education*, 9(1), 60-70.
- Koehler, M. J., Shin, T. S., & Mishra, P. (2012). How do we measure TPACK? Let me count the ways. In R. N. Ronau, C. R. Rakes, & M. L. Niess (Eds.), *Educational technology, teacher knowledge, and classroom impact: A research handbook on frameworks and approaches* (pp. 16-31). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-60960-750-0.ch002>
- Koh, J. H. L., Chai, C. S., & Tay, L. Y. (2014). TPACK-in-Action: Unpacking the contextual influences of teachers' construction of technological pedagogical content knowledge (TPACK). *Computers & Education*, 78(1) 20-29. <https://doi.org/10.1016/j.compedu.2014.04.022>
- Koh, J. H. L., Chai, C. S., & Tsai, C. C. (2010). Examining the technological pedagogical content knowledge of Singapore pre-service teachers with a large-scale survey. *Journal of Computer Assisted Learning*, 26(6), 563-573. <https://doi.org/10.1111/j.1365-2729.2010.00372.x>
- Koh, J. H. L., Chai, C. S., & Tsai, C. C. (2014). Demographic factors, TPACK constructs, and teachers' perceptions of constructivist-oriented TPACK. *Journal of Educational Technology & Society*, 17(1), 185-196.

- Kopcha, T. J., Ottenbreit-Leftwich, A., Jung, J., & Baser, D. (2014). Examining the TPACK framework through the convergent and discriminant validity of two measures. *Computers & Education*, 78(1), 87-96. <https://doi.org/10.1016/j.compedu.2014.05.003>
- Kopcha, T. J., & Sullivan, H. (2007). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development*, 55(6), 627-646. <https://doi.org/10.1007/s11423-006-9011-8>
- 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. <https://doi.org/10.1007/s11251-008-9075-4>
- Lin, T. C., Tsai, C. C., Chai, C. S., & Lee, M. H. (2013). Identifying science teachers' perceptions of technological pedagogical and content knowledge (TPACK). *Journal of Science Education and Technology*, 22(3), 325-336. <https://doi.org/10.1007/s10956-012-9396-6>
- Mansour, N. (2009). Science teachers' beliefs and practices: Issues, implications, and research agenda. *International Journal of Environmental and Science Education*, 4(1), 25-48.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332. <https://doi.org/10.3102/00346543062003307>
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK) the development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, 42(2), 123-149. <https://doi.org/10.1080/15391523.2009.10782544>
- Shafer, W. E., Fukukawa, K., & Lee, G. M. (2007). Values and the perceived importance of ethics and social responsibility: The US versus China. *Journal of Business Ethics*, 70(3), 265-284. <https://doi.org/10.1007/s10551-006-9110-9>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. <https://doi.org/10.3102/0013189X015002004>
- So, H. J., & Kim, B. (2009). Learning about problem based learning: Student teachers integrating technology, pedagogy and content knowledge. *Australasian Journal of educational technology*, 25(1), 101-116. <https://doi.org/10.14742/ajet.1183>
- Tondeur, J., van Braak, J., Sang, G., Voogt, J., Fisser, P., & Ottenbreit-Leftwich, A. (2012). Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. *Computers & Education*, 59(1), 134-144. <https://doi.org/10.1016/j.compedu.2011.10.009>
- Trochim, W. M., & Donnelly, J. P. (2008). *Research methods knowledge base* (3rd ed.). Mason, OH: Atomic Dog Publishing.
- Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge – a review of the literature. *Journal of Computer Assisted Learning*, 29(2), 109-121. <https://doi.org/10.1111/j.1365-2729.2012.00487.x>

Corresponding author: Yangyi Qian, qianyy@scnu.edu.cn

Australasian Journal of Educational Technology © 2017.

Please cite as: Deng, F., Chai, C. S., So, H.-J., Qian, Y., & Chen, L. (2017). Examining the validity of the technological pedagogical content knowledge (TPACK) framework for preservice chemistry teachers. *Australasian Journal of Educational Technology*, 33(3), 1-14. <https://doi.org/10.14742/ajet.3508>