ISSN 1392-0340 (Print)
ISSN 2029-0551 (Online)
https://doi.org/10.15823/p.2022.148.10

Pedagogika / Pedagogy
2022, t. 148, Nr. 4, p. 184–206 / Vol. 148, No. 4, pp. 184–206, 2022

VYTAUTO DIDŽIOJO UNIVERSITETO ŠVIETIMO AKADEMIJA

HOTS-Oriented TPACK Survey Validation

Siti Mina Tamah[1], Anita Lie[2], Imelda Gozali[3], Lusia Permata Sari Hartanti[4]

[1] Widya Mandala Surabaya Catholic University, 37 Kalijudan St., Surabaya, East Java, Indonesia, mina@ukwms.ac.id
[2] Widya Mandala Surabaya Catholic University, 37 Kalijudan St., Surabaya, East Java, Indonesia, anita@ukwms.ac.id
[3] Widya Mandala Surabaya Catholic University, 37 Kalijudan St., Surabaya, East Java, Indonesia, imelda.gozali@ukwms.ac.id
[4] Widya Mandala Surabaya Catholic University, 37 Kalijudan St., Surabaya, East Java, Indonesia, lusia.hartanti@ukwms.ac.id

Annotation. This study aimed to validate a TPACK scale which was Higher Order Thinking Skills (HOTS)-oriented, engaging 145 in-service teachers in assessing their self-perceived technological competence in teaching. The AMOS application was used for the Structural Equation Modeling (SEM) analysis and the Confirmatory Factor Analysis (CFA) computation. The results demonstrated that the survey questions were valid and reliable (five validity indicators categorized as 'fit' and CFA Construct Reliability = 0.85).

Keywords: *HOTS, language teachers, survey development, TPACK, validation.*

Introduction

As one of the models used to designate teachers' competencies for successful teaching incorporating technology, Technological Pedagogical Content Knowledge (TPACK) is commonly measured through survey questions. These survey items are one of the most widely used among the TPACK community (Koehler et al., 2013; Schmid et al., 2020).

Elaborated TPACK survey items were initially developed from items limited to PACK or PCK. With the absence of 'T' (Technological) knowledge, the PACK items were developed by Jang et al. (2009), following Shulman's model – comprising 28 items grouped into 4 constructs: namely Instructional Objects and Context (IOC), Instructional Representation and Strategies (IRS), Knowledge of Students' Understanding (KSU), and Subject Matter Knowledge (SMK). The items addressed students' perceptions of their college teachers' knowledge of pedagogy and content, or subject matter.

Other surveys to address teachers' knowledge of the global requirement to replace traditional instructional methods have also been developed. Similar to Jang et al. (2009), who employ a 5-point Likert scale, Sahin (2011), Akman and Guven (2015), and Schmid et al. (2020) have each developed a scale to study the technological, pedagogical, and content knowledge of teachers to include the 'T'element which does not appear in Jang et al.'s (2009). Sahin (2011) administered the survey to pre-service teachers majoring in English language education. Akman and Guven (2015) engaged social science teachers and teacher candidates, while Schmid et al. (2020) applied the survey to pre-service secondary school teachers.

To date, merely a few studies have been carried out to measure the validity and reliability of TPACK instruments in the Indonesian context. One of the few studies includes Zaeni et al.'s (2021) which involved pre-service teachers of mathematics. More studies are worth performing to provide more information on the instructional knowledge of teachers.

In preparing the students to sail the waters of the impending Society 5.0 era, teachers will need to possess analytical thinking skills in order to impart them to the students. In the field of education, analytical thinking is commonly termed as Higher-Order Thinking Skills (HOTS). However, recent studies demonstrated that teachers still grappled with the concept and application of HOTS in the classroom (Gozali et al., 2021; Lie et al., 2020).

Employing a valid and reliable survey to scrutinize teachers' knowledge and skills is partly an attempt to understand teachers' ongoing professional development which is an inevitable issue for successful teaching for the sake of students' quality learning. This paper then aims to shed light on the construct to develop a survey of technological, pedagogical, and content knowledge (TPACK) which incorporates Krathwohl's (2002) Higher Order Thinking Skills in the Apply, Analyze, Evaluate, and Create levels. The survey is specifically designed for language teachers

Literature Review

Technological Pedagogical Content Knowledge (TPACK)

In a teacher training institution, the knowledge of content (the C element in TPACK) of the preservice teachers is naturally given prime importance, especially at the beginning of the course. The curriculum in a language teaching institution, for instance, consists of some content subjects like the typical Listening, Speaking, Reading, and Writing (either discrete Listening subject, Speaking subject, Reading subject, and Writing subject, or integrated Listening & Speaking subject, and Reading & Writing subject).

Afterward, the P element in TPACK becomes as essential as the C element. In teaching method literature, the pedagogy-content knowledge model of Shulman (1986a) should not be neglected. Shulman (1986a; 1986b) argues that, for an effective and successful instructional experience, teachers should possess not only content or subject matter knowledge, but also the appropriate pedagogical methods or strategies to bring about the students' mastery of the content.

With the advancement of technology, the curriculum has been identified with the T element in TPACK. The typical course subject – with different names like CALL (Computer-Assisted Language Learning), IT in Language Education (courses offered in the researchers' teacher institution), Technology-Enriched Teaching, and Digital Class Management (courses offered in some teacher institutions) – is provided to focus on the key concepts to effectively integrate technology into teaching. This illustration shall provide how the PACK or originally PCK has transformed into TPACK.

The T element reveals the ICT integration in teaching. Teachers with high TPACK scores were found to be more adept at incorporating interactive media in their teaching modules (Fabian et al., 2019). Claims of great effectiveness for both teachers and students due to the T element have been reported in some studies (Ghavifekr & Rosdy, 2015; Mafang'ha, 2016; Stosic, 2015). Despite the barriers hindering teachers such as lack of support in technical matters and inadequate technology infrastructure (Akman & Guven, 2015; Syamdianita & Cahyono, 2021), teachers see the advantages of educational technology for their attitude, confidence, and competence. It is furthermore recommended that Information and Communications Technology (ICT) be used to support reflective teachers' continuing professional development (Lie et al., 2019; Lie et al., 2020). Schmidt et al. (2009) have earlier argued likewise that the use of TPACK as a framework for measuring teaching knowledge will influence the type and the design of pre-service teachers' training and in-service teachers' professional development programs.

TPACK is assumed to be an amalgam of the three foremost domains of knowledge of teachers. The TK, PK, and CK domains have developed into the TPACK framework through the existence of PCK, TCK, and TPK. Eventually, the resulting TPACK framework is formed through the interplay of the three augmented models, as illustrated in Figure 1. The definition of each of the TPACK subdomains is given in Table 1.

Figure 1

TPACK Framework (Koehler & Mishra, 2008 in Schmid et al., 2020)

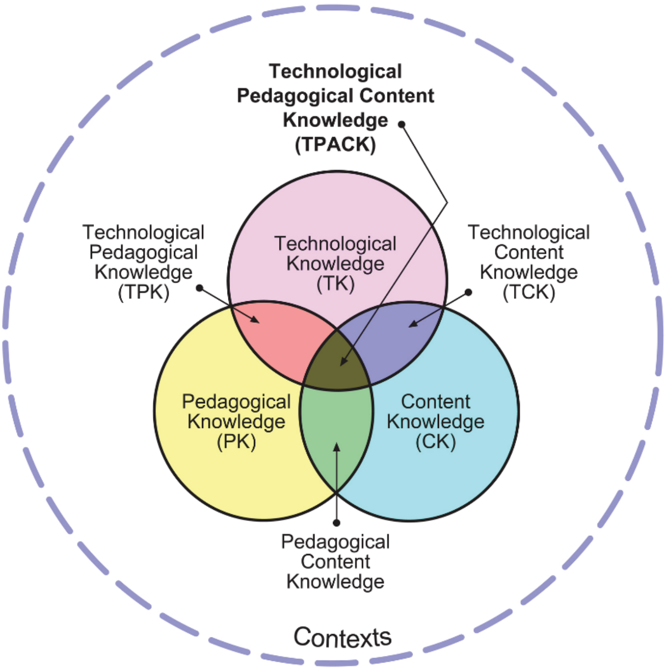


Table 1

The Definition of Each of the TPACK Subdomains

No.	Domain	Definition
1	Pedagogical Knowledge (PK)	Knowledge about process and practices or methods of teaching and learning and how it encompasses educational purposes, values, and aims (e.g. student learning, classroom management, lesson plan development, and implementation).
2	Content Knowledge (CK)	Knowledge about the actual subject matter that is to be taught (e.g., central facts, concepts, theories, procedures).
3	Technological Knowledge (TK)	Knowledge about standard technologies and how to operate them (e.g., from books and chalkboards to the internet and digital video).
4	Pedagogical Content Knowledge (PCK)	Knowledge about pedagogy that is applicable to the specific teaching content (e.g., knowing what teaching approaches fit the content, knowing how elements of content can be arranged for better teaching).
5	Technological Pedagogical Knowledge (TPK)	Knowledge of how teaching may be changed as the result of using particular technologies (e.g., knowing that a range of tools exists, the ability to select based on fitness, and knowledge of affordances of these tools for pedagogical practice).

No.	Domain	Definition
6	Technological Content Knowledge (TCK)	Knowledge about how technology and content are reciprocally related (e.g., knowing how subject matter can be changed by the application of technology).
7	Technological Pedagogical Content Knowledge (TPCK)	Knowledge for good teaching with technology which requires understanding how technologies can support teaching subject matter (e.g., knowing how technologies can help overcome problems in the process of teaching and learning, and how they can be used for constructive content and pedagogy).

Higher Order Thinking Skills

One simple definition of Higher order thinking skills or HOTS is "an ability in using and processing thought processes over the facts" (Lie et al., 2020, p. 2). Someone who has high-level thinking skills not only knows a certain fact but also uses the acquired knowledge to develop knowledge itself.

When it was revealed that the critical thinking ability of students in Indonesia is very low as measured in the Program for International Student Assessment (PISA), there were quite a lot of efforts to improve it. Some researchers have tried to study teachers' competence in asking HOTS-inspired questions in class (Gozali et al., 2021; Lie et al., 2020). Some others indicate the training provided to teachers (Barak & Judy 2009; Thompson, 2008; Gozali et al., 2021).

Conceptually, when HOTS is referred to, most educators prevalently refer to the learning taxonomies of Bloom, Anderson, and Krathwohl (Anderson & Krathwohl, 2001). These taxonomies are popular among practitioners for they tier thinking skills that are described clearly and easily understood.

Benjamin Bloom first conceived the taxonomy as comprising six cognitive categories, namely Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation, in the order of the lowest thinking level to the highest. Posteriorly, Anderson and Krathwohl refined the taxonomy further by adding a knowledge dimension (made up of factual, conceptual, procedural, and metacognitive knowledge types) (Anderson & Krathwohl, 2001; Miri et al., 2007) to Bloom's six-tiered cognitive processes. The cognitive processes themselves were then renamed to verb forms, and the "evaluation" and "synthesis" categories were exchanged. The resulting revised Bloom's Taxonomy then became Remember, Understand, Apply, Analyze, Evaluate, and Create, from the simplest to the most complicated (Krathwohl, 2002). Thus, HOTS is constituted by Analyse, Evaluate, and Create, the top three in the cognitive rungs.

TPACK and HOTS

The concurrent demand made to teachers to integrate technology into their teaching and to instil critical thinking capability into their students has prompted scholars to investigate the possible relationship between the two constructs, namely TPACK and HOTS. In general, technological integration in the classroom is seen as an enabler of students' critical thinking and problem-solving skills (Pasani, 2018). Hence, TPACK is conceived as a teaching approach that is able to promote HOTS ability in students (Susilawati & Khaira, 2021). Researchers have also developed various TPACK-inspired teaching media (module, virtual laboratory, textbook) designed to foster HOTS (Bakri et al., 2021; Ilmi & Sunarno, 2020; Zainuddin et al., 2021).

In light of the above phenomena, the TPACK questionnaire has been further developed to include other components such as 21st-century skills, hence producing the TPACK21 scale (Valtonen et al., 2015; 2018). Zaeni et al. (2021) have also developed a HOTS-based TPACK questionnaire to assess the self-perception of mathematics preservice teachers in this respect. However, Zaeni et al.'s work, as well as the other research investigating the interplay between TPACK and HOTS, is situated within the context of hard sciences. The dearth of studies in the social sciences concerning the integration of TPACK and HOTS necessitates the development and validation of a HOTS-infused TPACK scale which will be able to advance quantitative investigation in this field.

Research Methodology

Sample

The validation of the survey questionnaire engaged a total of 145 teachers of English and Bahasa Indonesia from 50 senior high schools, 35 junior high schools, and 5 elementary schools in Indonesia. They were from 56 cities in 13 provinces in Indonesia. Forty-eight (33%) teachers were male and 97 (67%) were female. As for teaching experience, 16 (11.03%) reported having below 5 years of experience (novice); 45 (31.03%) between 5–10 years (apprentice); 49 (33.79%) between 10–15 years (practitioner), and 35 (24.14%) more than 15 years (senior).

Data Collection Procedure

The very first step in the data collection procedure was the researchers' attempt to meet the ethical standards of doing educational research with human subjects. The participants filled out a consent form to indicate their agreement to participate in this study in a voluntary manner, and their awareness of their rights as study participants. They were assured about the anonymity and confidentiality of their data and use for the sole purpose of the study. They were also informed of the general aim of this study.

While waiting for the consent form completion, the other step involved the formulation of survey items – methodically termed item pooling. Two main references taken from Rolando et al. (2021) and Zaeni et al. (2021) guided the researchers to formulate the survey items. Although the participants were English and Indonesian teachers, the survey items were prepared in one version – Indonesian only. The survey comprised 26 items classified into seven categories as summarized in Table 2.

Table 2

Number of TPACK Items in Each Category

Categories	Number of Items
Pedagogical Knowledge	5
Content Knowledge	3
Technological Knowledge	3
Technological – Pedagogical Knowledge	4
Technological – Content Knowledge	3
Pedagogical – Content Knowledge	3
Technological – Pedagogical – Content Knowledge	5
Total	26

All 26 statements contain one or a combination of the verbs denoting higher-order thinking skills in Krathwohl's framework (2002): Apply, Analyze, Evaluate, and Create (See the Appendix for a complete listing of the statements and their HOTS leveling). The typical 'have' as in the original survey 'I have sufficient knowledge about my teaching subject' (as in Schmid et al., 2020), 'I have knowledge in solving a technical problem with the computer' (as in Sahin, 2011), and 'I have the technical skills to use computers effectively' (as in Rolando et al., 2021) was reworded to incorporate the HOTS expectations. Further combining and rewording had been performed by the researchers. From Zaeni et al.'s TK Technological Knowledge (TK) items, 2 of the 6 items were combined and reworded. The same was done for 2 of the 5 items in Rolando et al.'s Technological Knowledge (TK). They became 1 item reformulated as 'I can learn and use new technology easily.'

The item wording to incorporate the HOTS element was emphasized in the TPK category. The items were not meant for teachers only; the students' interests were also taken into account. As an example, the item which typically appeared as 'I can choose an appropriate teaching method to solve students' problems in mastering the subject matter.' and 'I am able to facilitate my students to use technology to find more information on their own.' as used in Zaeni et al. (2021) and Rolando et al. (2021) respectively have been reformulated as follows: 'I can use technology to help students find information on their own.' (See item 2 in the TPK category in the Appendix).

This item pool implicitly indicated that expert judgments had been realized by examining the relevant literature, reading and rereading the existing items which became the primary references. The principal intention was to come closer to the definition followed earlier. The researchers maintained that teachers' understanding of how to use technology tools was connected with TK, while TCK was related to teachers' ability in making use of technology to present the content. Consequently, the TK items in this study amounted to only 3 items instead of 5 (Rolando et al., 2021) and 6 (Zaeni et al., 2021).

A pilot test was administered to five language teachers in a school in order to examine their understanding of the survey questions and time estimation for the survey completion. The link to the instruments, produced in an online format via Google Forms, was sent via WhatsApp Groups to the potential participants. These WhatsApp Groups consisted of alumni from the teacher certification cohorts.

Ensuring the validity of the survey, the three researchers read TPACK literature and got together to establish the conceptual framework for the instrument. Items were reread upon feedback from the tryout group. As presented previously, altogether 26 items were generated (See Table 2 above). The survey used a Likert scale to provide the four choices of strongly disagree (1), disagree (2), agree (3), and strongly agree (4).

Data Analysis Procedure

Confirmatory Factor Analysis (CFA) is, as claimed by Malkanthie (2015), appropriate when researchers have some understanding through theory, empirical research, or both of the latent variable structures. Chi-square functions as an analytical tool is used to examine or validate the accuracy of the proposed theoretical structure. This Chi-square statistical tool is very sensitive to sample size (a small sample size is harder to bring about accurate statistical significance), so caution is needed in drawing conclusions based solely on the significance of the chi-square test. Byrne (2016) points out that it would be better and more correct to make a decision based on other fit indices such as CFI, RMSEA, and SRMR. As an alternative, Iacobucci (2010) suggests the use of a ratio of the chi-square test statistic to the degrees of freedom ($\chi 2/df$) of which its cut-off value for the ratio is ≤ 3 to result in a fit category. Meanwhile, Byrne (2016) argues slightly differently to keep the cut-off value of ≤ 2 for $\chi 2/df$.

Keeping the theoretical issue in mind with regard to data analysis, the researchers would check the obtained data for the validity of the survey items by employing Confirmatory Factor Analysis (CFA) instead of Exploratory Factor Analysis (EFA). Prior to CFA, checking the multivariate normality was carried out.

Using structural equation modeling (SEM) maintaining the Amos Graphical approach, this study would report six Amos outputs for the validity measurement: Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square of Approximation (RMSEA), Goodness of Fit Index (GFI), Standardized Root Mean Square Residual (SRMR), and CMIN/df. In brief, this study made use of the Structural Equation Modeling (SEM) analysis approach using AMOS software to quantitatively analyze the data.

To check the reliability of the survey items, the researchers did not employ Cronbach's Alpha. Instead, Construct Reliability (CR) was used. CR is, as defined by Hair et al. (2019), a measure of the internal consistency and reliability of measured variables that describe a latent construct. Hair et al. (2019) further assert that a construct which has the Construct Reliability (CR) value of 0.70 or greater indicates that the items are reliable.

Results and Discussion

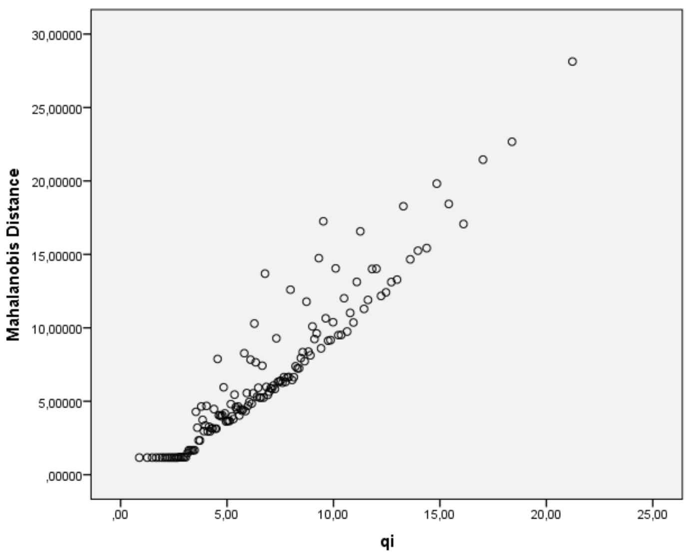
Normally Distributed Multivariate Check

Prior to Confirmatory Factor Analysis (CFA) statistical analysis, the data were, as previously mentioned, checked for their Maximum Likelihood (ML) to see if the data were normally distributed multivariate. One way to check multivariate normality is, as Ramzan et al. (2013) point out, by the use of a Q-Q plot Mahalanobis Distance.

In this study, software SPSS 22 was employed to obtain the Q-Q plot Mahalanobis Distance, the result of which is presented in Figure 2. It was revealed that the Q-Q plot indicated a linear pattern. The implication is that the multivariate data were normally distributed.

Figure 2

Result of the Q-Q Plot Mahalanobis Distance



Statistical analysis was performed by checking Pearson's correlation coefficient. One requirement to keep in mind is that when Pearson's correlation coefficient is greater than the table percent point of the normal probability plot correlation coefficient, the H0 is rejected which means that the data are not distributed normally and vice versa. The H0 is accepted when the Pearson's correlation coefficient is smaller than the table percent point.

As seen in Table 3, the Pearson's correlation coefficient was found to be 0.947. This implies that the Pearson's correlation coefficient was smaller than the table percent point (t-table > 0.988; n=145; α =0.05); therefore, the conclusion is that multivariate normality was confirmed.

Table 3

Result of Pearson's Correlation Coefficient Analysis

		Mahalanobis Distance	**qi**
Mahalanobis Distance	Pearson Correlation	1	.947**
	Sig. (2-tailed)		.000
	N	145	145
qi	Pearson Correlation	.947**	1
	Sig. (2-tailed)	.000	
	N	145	145

**. Correlation is significant at the 0.01 level (2-tailed).

Confirmatory Factor Analysis (CFA)

Confirmatory Factor Analysis (CFA) is used to confirm the construct validity of any developed instrument. Furthermore, as Confirmatory Factor Analysis (CFA) is appropriate when researchers have some understanding through theory, empirical research, or both of the latent variable structures as claimed by Malkanthie (2015), this study used CFA. It was utilized to measure observed and latent variables to describe or represent a number of a factor.

A parameter estimate test of Maximum Likelihood (ML) was performed after the data were found to be normally distributed. Hair et al. (2019) put forward that in CFA, the test most prevalently used with regard to goodness of fit includes the following indicators: Comparative Fit Index (CFI), Root Mean Square Error of approximation (RMSEA), Goodness-of-Fit (GFI), and Tucker Lewis Index (TLI). The goodness of fit is used to examine whether the model is theoretically similar to the empirical model. This corresponds to the claim by Chua (2012) pointing out similarly that most research employs Chi-Square, CFI, TLI, GFI, and RMSEA. Hair et al. (2019) further contend that the inclusion of four or five components in the goodness of fit criteria is adequate to examine the appropriateness of a model – after ensuring the absolute fit indices, incremental fit indices, and parsimonious fit. Table 4 shows the result of testing the goodness of fit, which in turn reveals that the CFA results using AMOS for the hypothesized seven-factor (PK, CK, TK, PCK, TPK, TCK, TPACK) model are excellent.

Table 4

Test Result of Goodness of Fit

No.	Criteria	Test Result of Fit Value	Cut-off Value	Conclusion
1	*Comparative Fit Index* (CFI)	0.912	≥ 0.9	*Fit*
2	*Tucker Lewis Index* (TLI)	0.9	≥ 0.9	*Fit*
3	*Root Mean Square of Approximation* (RMSEA)	0.076	0.03-0.08	*Fit*
4	*Goodness of Fit Index* (GFI)	0.8	≥ 0.9	*Marginal Fit*
5	*Standardized Root Mean Square Residual* (SRMR)	0.05	≤ 0.08	*Fit*
6	CMIN/df	1.833	≤ 2.0	*Fit*
	χ^2	509.639		

Comparative Fit Index (CFI)

Comparative Fit Index (CFI) measures the relative goodness-of-fit in comparison to a simpler model. A cutoff value greater than 0.9 (≥ 0.9) indicates the fit criteria (Chua, 2012; Hair et al., 2019). Waluyo (2016) affirms that the CFI value of 1 indicates the model with the highest fit. As the CFI value obtained from the data analyzed was 0.912, it can be concluded that a fit criterion had been obtained between the hypothesized model and the observed data. Therefore, the model could be accepted.

Tucker Lewis Index (TLI)

TLI is used to compare a measured model with a baseline model. A cutoff value greater than 0.9 (≥ 0.9) indicates the fit criteria for TLI (Chua, 2012; Hair et al., 2019). The TLI value obtained from the data analyzed was found to be 0.9 resulting in the confirmation that the model was accepted.

Root Mean Square Error of Approximation (RMSEA)

One most frequently used statistical measurements to accept or reject a model with a large sample or observed variables is the Root Mean Square Error of Approximation (RMSEA) (Hair et al., 2019). In this study, the RMSEA obtained value was 0.076 (a value between 0.03-0.08). Therefore, the model fulfilled the fit criteria.

Goodness of Fit Index (GFI)

Goodness of Fit examines the compatibility of the observed or the actual input (covariance matrix or correlation) and prediction of the proposed model. Chua (2012) and Hair et al. (2019) argue that the cutoff value for GFI is ≥ 0.9. The value of GFI (≥ 0.90) indicates a good fit, while the value between 0.8–0.9 (0.80 ≤ GFI ≤ 0.90) is identified as a marginal fit. Table 4 shows that the obtained value of GFI was 0.8. This indicates that in this study the model belonged to marginal fit.

Standardized Root Mean Square Residual (SMRM)

SRMR shows the extent to which a model prediction matches the data perfectly. The higher the value, the worse it is (Kyndt & Onghena, 2014). SRMR depends tremendously on factor loadings to measure the model and it is relatively less sensitive to violations of distributional assumptions. SMRM of below 0.8 (the cutoff value) is required for a model to be accepted (Hu & Bentler, 1999; Kyndt & Onghena, 2014). The statistical analysis for SMRM presented in Table 4 indicates that SMRM was reported to be 0.05; therefore, the model was accepted.

Chi-square and the Ratio of the Chi-Square Test Statistic to the Degrees of Freedom (CMIN/df)

Table 4 presents the statistical result for $\chi2/df$ analysis. The value of 1.833 was reported indicating that the model belonged to the fit category. The goodness of fit parameter in Table 4 shows that all the criteria generated by the model met the fit goodness.

Further statistical analysis is performed to look into the factor loadings. The factor loadings of each criterion or indicator are presented in Table 5 and the model is shown in Figure 3. Table 5 summarizes the results of calculating the loading factor of each item survey instrument. All indicators in the CFA had positive and significant loadings, ranging from 0.585 to 0.909, signifying the strong weight of each item in the model. There are no items with a loading factor of less than 0.05, so no items are omitted from the model. Five sub-indicators measured the PK indicator with factor loading ranging from 0.83-0.909. Considering the CK indicator, the sub-indicators with the highest factor loading were CK2 (0.878), which means that CK2 (I can **solve** problems related to English) is the sub-indicator that contributes significantly to the CK construct. For the TK indicator, the sub-indicator with the highest factor loading was TK2 (0.842), indicating that the sub-indicator that largely contributes to the TK construct is TK2 (I can **solve** my technical problems using technology). For the PCK indicator, the sub-indicator having the highest factor loading was PCK2 (0.867), which means PCK2 (Without technology, I can choose a suitable learning method for learning English.) is the sub-indicator that contributes significantly to the CK construct. Considering the TPK indicator, the sub-indicators having the highest factor loading was TPK3 (0.856), which means TPK3 (Make students **design** forms of information representation in various ways (text, graphics, videos, comics, etc.) is the sub-indicator that contributes significantly to the TPK construct. For the TCK indicator, the sub-indicator having the highest factor loading was TCK1 (0.823), indicating that the sub-indicator that largely contributes to the TCK construct is TCK1 (I can **search and use** technology created specifically for English). For the TPACK indicator, the sub-indicator having the highest factor loading was TPACK5 (0.88), indicating that the sub-indicator that largely contributes to the TPACK construct is TPACK5 (I can **help** my colleagues integrate technology, pedagogy, and content in my school).

Hair et al. (2019) state that factor loadings ranging from ± 0.30 to ± 0.40 are considered to meet the minimum level for structural interpretation, while factor loadings of ± 0.50 or more are considered significant. It can be seen from Table 5 that all of the factor loadings values are greater than 0.50, which means that they can be deemed significant.

All items have exceeded the minimum value of loading factors > 0.50, making them valid in construct or acceptable. Figure 3 shows the finalized measurement model and the result of the calculation of the CFA procedures, which functions to determine the extent of the relationship between sub-indicators with indicators, as can be seen from the value of loading factors in the estimated standard. The finalized model (Figure 3) became the baseline model for the next analyses.

Table 5

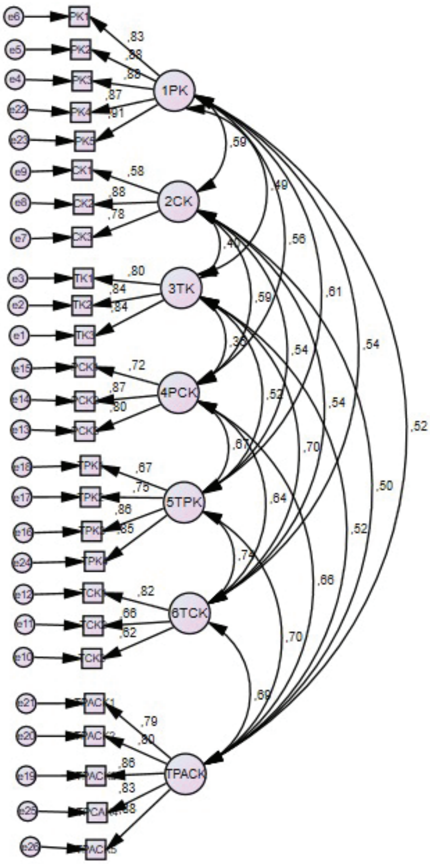
Factor Loadings

No.	Major Indicator	Sub-Indicators		Factor Loadings
1	1PK	1	PK1	0.83
		2	PK2	0.878
		3	PK3	0.875
		4	PK4	0.865
		5	PK5	0.909
2	2CK	6	CK1	0.585
		7	CK2	0.878
		8	CK3	0.776
3	3TK	9	TK1	0.795
		10	TK2	0.842
		11	TK3	0.84
4	4PCK	12	PCK1	0.718
		13	PCK2	0.867
		14	PCK3	0.8

No.	Major Indicator	Sub-Indicators		Factor Loadings
5	5TPK	15	TPK1	0.671
		16	TPK2	0.754
		17	TPK3	0.856
		18	TPK4	0.846
6	6TCK	19	TCK1	0.823
		30	TCK2	0.661
		31	TCK3	0.615
7	TPACK	22	TPACK1	0.788
		23	TPACK2	0.804
		24	TPACK3	0.857
		25	TPACK4	0.827
		26	TPACK5	0.88

Figure 3

TPACK Confirmatory Factor Analysis



Construct Reliability (CR)

Cronbach's Alpha is the most widely used approach for determining reliability and internal consistency. Other options, such as Construct Reliability (CR), have also been accepted. CFA may therefore be used to examine not only construct validity but also construct reliability (Hair et al., 2019). Hair et al. (2019) further assert that a construct which has the Construct Reliability (CR) value of 0.70 or greater (CR ≥ 0.70) belongs to the reliable category.

Unlike the previous related studies, this study did not employ Cronbach's statistical analysis. The research data having been analyzed (Table 6) resulted in CFA Construct Reliability values ranging from 0.71 to 0.94 (all values were greater than 0.7). As a consequence, the reliability of the survey items was confirmed. This indicates that the developed instruments meet the requirements for reliability and can be relied upon to offer information about the HOTS-Oriented TPACK survey.

Table 6
CFA Construct Reliability

No.	Indicator	Factor Loadings	Error	Construct Reliability (CR)
1	PK1	0.83	0.187	
2	PK2	0.878	0.244	
3	PK3	0.875	0.234	0.94
4	PK4	0.865	0.246	
5	PK5	0.909	0.196	
6	CK1	0.585	0.753	
7	CK2	0.878	0.192	0.8
8	CK3	0.776	0.307	
9	TK1	0.795	0.396	
10	TK2	0.842	0.321	0.86
11	3TK3	0.84	0.321	
12	PCK1	0.718	0.488	
13	PCK2	0.867	0.24	0.85
14	PCK3	0.8	0.313	
15	TPK1	0.671	0.544	
16	TPK2	0.754	0.401	0.86
17	TPK3	0.856	0.323	
18	TPK4	0.846	0.313	

19	TCK1	0.823	0.394	
20	TCK2	0.661	0.759	0.71
21	TCK3	0.615	0.645	
22	TPACK1	0.788	0.434	
23	TPACK2	0.804	0.366	
24	TPACK3	0.857	0.132	0.92
25	TPACK4	0.827	0.367	
26	TPACK5	0.88	0.126	

As presented earlier, the data analysis of the structural equation modeling (SEM) using Amos in this study has indicated the valid measurement of the survey items. Five of the six indicators have been reported as 'fit'; with only one 'marginal fit'. Comparative Fit Index (CFI) is reported to be 0.912, Tucker Lewis Index (TLI) 0.9, Root Mean Square of Approximation (RMSEA) 0.076, Standardized Root Mean Square Residual (SRMR) 0.05, and CMIN/df 1.833. These all have been validated as 'fit' measurements. Only a Goodness of Fit Index (GFI) of 0.8 is reported to be 'marginal fit'. Once the CFA computation was completed, the instrument of the HOTS-Oriented TPACK model was confirmed in 26 sub-indicators, divided into seven indicators (see Appendix).

In light of the reliability measurement, the CFA Construct Reliability value has been reported to range between 0.71 and 0.94 indicating the confirmation of the reliable survey items which are HOTS-oriented. This study has particularly substantiated Zaeni et al. 's (2021) study which shows a reliability range between 0.59 and 0.77. When compared with theirs, it is seen that the values reported here have higher reliability. Yet, further studies need to be conducted as the reliability measurements in Zaeni et al.'s (2021) study and this study are not the same. Zaeni et al.'s (2021) study used Cronbach's while this study uses Construct reliability. More studies can be conducted although there is an argument that analyses employing Construct reliability and Cronbach's will yield a similar result.

Rolando's 29 item-survey, which measures on a 7-point Likert scale and has been proven to be valid, is used in Rolando et al.'s (2021) study. This study's finding has simplified the Likert scale by the use of a 4-point Likert scale with fewer items (26 items). The simplification from a 7-point Likert scale to a 4-point Likert scale is done to avoid the 'middle' point of '4' which means that the use of a 6-point Likert scale can be used. Yet, the use of a 4-point Likert scale, as Chang (1994) argues, has higher reliability than the 6-point scale. Therefore, the researchers chose to keep a more robust, 4-point Likert scale. More studies are not unwise to be carried out to get more conclusive information.

Revealed as one of its weaknesses, this study does not include validation of the translated version of the survey items for its Indonesian teacher sample. Further studies can consequently be performed to follow Sahin (2011) who has provided a step-by-step model for the procedure of translating items in a survey. Another limitation is perhaps the lack of heterogeneity among the participants; they were mostly teachers who had graduated from a Teacher Professional Development program which was conducted online, and so the teachers were mostly familiar with technology. Thus, higher inclusivity and a greater number of participants, like in Inceoglu and Aslan (2022) would have made the results more generalizable.

Conclusion

The study has tried to develop a set of self-perceived survey items to measure teachers' knowledge of Technology, Pedagogy, and Content incorporating Higher Order Thinking Skills (HOTS) in teaching languages. The development of the items has been validated and this study has brought about a valid and reliable HOTS-oriented TPACK instrument. Despite its limitation, as revealed in the Discussion section, this study has contributed to the TPACK literature. Particularly, this study has resulted in a valid and reliable TPACK survey instrument which incorporated HOTS-denoting verbs. The items have been reformulated using words such as "search and use", "explain", "choose and solve", and "develop, organize, and collaborate" indicating "apply", "analyze", "evaluate" and "create" domains respectively. Therefore, the study result can be used to uncover language teachers' perspectives when technology-infused learning environments are implemented, and the instrument can be used in other studies.

Acknowledgment

The research has been funded by Widya Mandala Surabaya University Research Institute. The writers are therefore grateful to their colleague, Dr. Hartono Pranjoto who chairs the Research Institute.

References

Akman, O., & Guven, C. (2015). TPACK survey development study for social sciences teachers and teacher candidates. *International Journal of Research in Education and Science* (IJRES), *1*(1), 1–10.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing.* Longman.

Bakri, F., Kusuma, R., & Permana, A. H. (2021, October). TPACK and augmented reality in kinematics practicum module: forming hots physics education students. In *Journal of Physics: Conference Series*, *1*, 012041. IOP Publishing.

Byrne, B. M. (2016). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (3rd ed.). Routledge Taylor & Francis Group.

Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, *18*, 205–215.

Chua, Y. P. (2012). *Mastering research methods*. Mcgraw-Hill Education.

Ghavifekr, S. & Rosdy, W.A.W. (2015). Teaching and learning with technology: Effectiveness of ICT integration in schools. *International Journal of Research in Education and Science* (IJRES), *1*(2), 175–191.

Gozali, I., Lie, A., Tamah, S. M., & Jemadi, F. (2021). HOTS questioning ability and HOTS perception of language teachers in Indonesia. *Indonesian Journal of Applied Linguistics*, *11*(1), 60–71. https://doi.org/10.17509/ijal.v11i1.34583

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Black, W. C., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning EMEA. https://doi.org/10.1002/9781119409137.ch4

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003

Ilmi, A. M., & Sunarno, W. (2020). Development of TPACK based-physics learning media to improve HOTS and scientific attitude. *Journal of Physics: Conference Series*, *1440* (1), 012049. IOP Publishing.

İnceoğlu, S. Ö., & Aslan, D. (2022). A study of the validity and reliability of the Preschool School Climate Scale (PSCS). *Pedagogika, 145*(1), 22–41.

Jang, S. J., Guan, S. Y., & Hsieh, H. F. (2009). Developing an instrument for assessing college students' perceptions of teachers' pedagogical content knowledge. *Procedia Social and Behavioral Sciences*, *1*, 596–606. https://doi:10.1016/j.sbspro.2009.01.107

Koehler, M. J., Mishra, P., & Cain, W. (2013). What is Technological Pedagogical Content Knowledge (TPACK)? *Journal of Education, 193*(3), 13–19.

Krathwohl, D. R. (2002). A revision of bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kyndt, E., & Onghena, P. (2014). The Integration of work and learning: Tackling the complexity with Structural equation modelling. *Professional and Practice-based Learning*, *9*, 255–291. https://doi.org/10.1007/978-94-007-7012-6_14

Lie, A., Tamah, S. M., Trianawaty, & Jemadi, F. (2019). Challenges and resources in enhancing English teachers' proficiency. *The GLOCAL Proceedings Asia 2019 The GLOCAL in Asia 2019-The GLOCAL CALA 2019,* 306–317. https://glocal.soas.ac.uk/cala2019-proceedings/

Lie, A., Tamah, S. M., Gozali, I., & Triwidayati, K. R. (2020). *Mengembangkan Keterampilan Berpikir Tingkat Tinggi* [*Developing the skills of higher order thinking*]. PT Kanisius.

Lie, A., Tamah, S. M., Gozali, I., Triwidayati, K. R., Utami, T. S. D., & Jemadi, F. (2020). Secondary school language teachers' online learning engagement during the COVID-19 pandemic in Indonesia. *Journal of Information Technology Education: Research* (JITE-R), *19*, 803–832. https://doi.org/10.28945/462610.28945/4626

Mafang'ha, M. (2016). Teachers' experience on the use of ICT to facilitate teaching: A case of Ilala district secondary schools [Doctoral dissertation. University of Tanzania]. http://repository.out.ac.tz/1584/1/Mafanga-Dissertation-14-10-2016-Final.pdf

Malkanthie, A. (2015). *Structural equation modeling with AMOS*. Nippon Graphics Pvt. Ltd. https://www.researchgate.net/ publication/278889068

Mintre, R., & Lie, A. (2020). Higher order thinking questions in teaching reading in a private primary school. *Beyond Words*, *8*(1), 36–46. https://doi.org/10.33508/BW.V8I1.2092

Miri, B., David, B. C., & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: A case of critical thinking. *Research in Science Education*, *37*(4), 353–369. https://doi.org/10.1007/s11165-006-9029-2

Pasani, C. F. (2018, August 4). *Tpack untuk mengembangkan HOTS dan berbagai literasi. [TPACK to develop HOTS and other forms of literacies]* [Paper presentation]. *Proceedings of Seminar Nasional Pendidikan Matematika* I, Banjarmasin, South Kalimantan, Indonesia. http://eprints.ulm.ac.id/4007/1/1.pdf.

Ramzan, S., Zahid, F. M., & Ramzan, S. (2013). Evaluating multivariate normality: A graphical approach. *Middle East Journal of Scientific Research*, *13*(2), 254–263. https://doi.org/10.5829/idosi.mejsr.2013.13.2.1746

Rolando, L. G. R., Salvador, D. F., Vasconcellos, R. F. R. R., & Luz, M. R. M. P. D. (2021). TPACK for meaningful learning survey: «Paths» for professional development of biology teachers in Brazil. *TOJET: The Turkish Online Journal of Educational Technology*, *20*(2), 168–181.

Sahin, I. (2011). Development of survey of technological pedagogical and content knowledge (TPACK). *TOJET: The Turkish Online Journal of Educational Technology*, *10*(1), 97–105.

Schmid, M., Brianza, E., & Petko, D. (2020). Developing a short assessment instrument for technological pedagogical content knowledge (TPACK.xs) and comparing the factor structure of an integrative and a transformative model. *Computers & Education, 157*, 103967. https://doi.org/10.1016/j.compedu.2020.103967

Schmidt, D. A., Baran, E., Thompson, A. D. Mishra, P., M. J. 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.

Shulman, L. S. (1986a). Paradigms and research programs in the study of teaching: A contemporary perspective. In Wittrock, M. C. (Ed.), *Handbook of research on teaching* (3rd ed.) (pp. 3–36). McMillan Publishing Company.

Shulman, L. S. (1986b). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Stosic, L. (2015). The importance of educational technology in teaching. *International Journal of Cognitive Research in Science, Engineering and Education*, *3*(1), 111–114.

Susilawati, E., & Khaira, I. (2021). Higher Order Thinking Skill (HOTS) dan model pembelajaran TPACK serta penerapannya pada matakuliah Strategi Pembelajaran PPKn. [HOTS and TPACK-based learning model and its application in PPKn course.] *Jurnal Teknologi Pendidikan (JTP)*, *14*(2), 139–147. https://doi.org/10.24114/jtp.v14i2.28338

Syamdianita, S., & Cahyono, B. Y. (2021). The EFL pre-service teachers' experiences and challenges in designing teaching materials using TPACK framework. *Studies in English Language and Education*, *8*(2), 561–577.

Valtonen, Teemu & Sointu, Erkko & Mäkitalo, Kati & Kukkonen, Jari. (2015). Developing a TPACK measurement instrument for 21st century pre-service teachers. *Seminar.net – International Journal of Media, Technology and Lifelong Learning*, *11*, 87–100. doi: 10.7577/seminar.2353

Valtonen, T., Sointu, E., Kukkonen, J., Kontkanen, S., Lambert, M. C., & Mäkitalo-Siegl, K. (2017). TPACK updated to measure pre-service teachers' twenty-first century skills. *Australasian Journal of Educational Technology*, *33*(3), 15–31. https://doi.org/10.14742/ajet.3518

Waluyo, M. (2016). Mudah cepat tepat penggunaan Tools Amos dalam aplikasi (SEM). *UPN Jatim*.

Zaeni, A., Rahayu, W, & Makmuri (2021). Pengembangan instrumen *self-assessment technological pedagogical content knowledge* (TPACK) calon guru matematika berbasis HOTS. *Teorema*: *Teori dan Riset Matematika*, *5*(2), 59–68. https://jurnal.unigal.ac.id/index.php/teorema/article/view/4960

Zainuddin, M., Kharis, M., Waluyo, B., & Nahdiyah, U. (2021). Integrating TPACK based HOTS-Textbooks: A case study to attest teaching style in primary school. *Review of International Geographical Education Online*, *11*(5), 3662–3670.

Appendix

The TPACK-HOTS Self-Report Survey

	TPACK Domains	HOTS Categories
A	**Technological Knowledge**	
1	I can learn and use new technology easily.	(C3 Apply)
2	I can **solve** my technical problems using technology.	(C5 Evaluate)
3	I can be **creative** using technology.	(C6 Create)
B	**Pedagogical Knowledge**	
1	I can choose **learning** strategies according to the needs and conditions of students.	(C5 Evaluate)
2	I can **manage/organize** the steps of the learning method to make it easier for students to understand the material.	(C6 Create)
3	I can **manage/organize** the class so that students don't get bored in learning.	(C6 Create)
4	I can **choose** the correct form of assessment according to the characteristics of the material.	(C5 Evaluate)
5	I can **choose** the form of assignments that help students to think critically.	(C5 Evaluate)
C	**Content Knowledge**	
1	I judge that my **knowledge** of teaching English is sufficient.	(C5 Evaluate)
2	I can **solve** problems related to English.	(C5 Evaluate)
3	I can **develop** a deeper understanding of English.	(C6 Create)
D	**Technological – Content Knowledge**	
1	I can **search and use** technology created specifically for English.	(C3 Apply)
2	I consider that my **knowledge** of technology for English research is sufficient.	(C5 Evaluate)
3	With technology, I can **collaborate** with colleagues to deepen my knowledge of English.	(C6 Create)
E	**Pedagogical-Content Knowledge**	
1	Without technology, I can **explain** various theories and problems in English science.	(C4 Analyze)
2	Without technology, I can **choose** a suitable learning method for learning English.	(C5 Evaluate)
3	Without technology, I can **arrange** the stages of material to support the understanding of English.	(C6 Create)

	TPACK Domains	HOTS Categories
F	**Technological-Pedagogical Knowledge: I can use technology to**	
1	Make students **apply** their knowledge in the real world.	(C3 in students)
2	Help students **find** information on their own.	(C5 in students)
3	Make students **design** forms of information representation in various ways (text, graphics, videos, comics, etc.).	(C6 in students)
4	Make students **collaborate** with each other in using technology.	(C6 in students)
G	**Technological-Pedagogical-Content Knowledge**	
1	I can **combine** technology with the methods used to teach English content.	(C6 Create)
2	I can **choose** technology in my classroom to improve what I teach, how I teach, and what students learn.	(C5 Evaluate)
3	I can **create** independent learning activities with technology for learning English.	(C6 Create)
4	I can **evaluate** English learning combined with technology based on indicators.	(C5 Evaluate)
5	I can **help** my colleagues integrate technology, pedagogy, and content in my school.	(C5 Evaluate)

Į aukštesnio lygio mąstymo įgūdžius orientuoto technologinių pedagoginių turinio žinių klausimyno validumas

Siti Mina Tamah[1], Anita Lie[2], Imelda Gozali[3], Lusia Permata Sari Hartanti[4]

[1] Vidjos Mandalos Surabajos katalikų universitetas, Kalijudan g. 37, Surabaja, Rytų Java, Indonezija, mina@ukwms.ac.id
[2] Vidjos Mandalos Surabajos katalikų universitetas, Kalijudan g. 37, Surabaja, Rytų Java, Indonezija, anita@ukwms.ac.id
[3] Vidjos Mandalos Surabajos katalikų universitetas, Kalijudan g. 37, Surabaja, Rytų Java, Indonezija, imelda.gozali@ukwms.ac.id
[4] Vidjos Mandalos Surabajos katalikų universitetas, Kalijudan g. 37, Surabaja, Rytų Java, Indonezija, lusia.hartanti@ukwms.ac.id

Santrauka

Technologinės pedagoginės turinio žinios (angl. *TPACK*) tampa vienu iš svarbiausių konstruktų, apibūdinančių mokytojų kompetencijas, susijusias su technologijomis grindžiamu mokymu. Šioms žinioms vertinti dažniausiai naudojamas apklausos klausimų rinkinys. Apklausos elementai paprastai formuluojami remiantis šiais septyniais konstruktais: technologinės žinios, pedagoginės žinios, turinio žinios, technologinio turinio žinios, pedagoginio turinio žinios, technologinės pedagoginės žinios ir technologinio pedagoginio turinio žinios. Atsižvelgiant į tai, kad mokytojai vienu metu integruoja technologijas ir ugdo mokinių kritinio mąstymo įgūdžius, šiuo tyrimu siekiama patvirtinti technologinių pedagoginių turinio žinių (angl. *TPACK*) priemonių rinkinį, kuris orientuotas į aukštesnio lygio mąstymo įgūdžius (angl. *HOTS*). Tyrime dalyvavo 145 mokytojai iš įvairių Indonezijos sričių.

Šis tyrimas taip pat skirtas įvertinti į aukštesnio lygio mąstymo įgūdžius orientuotus technologinių pedagoginių turinio žinių tyrimo elementus. Siekiant šio tikslo remiamasi klausimynu, kuris sudarytas iš 26 teiginių ir matuoja savivokos kompetencijas technologijų perpildytame mokyme. Tyrime taikytas struktūrinių lygčių modeliavimo (angl. *SEM*) analizės metodas naudojant AMOS programinę įrangą. Validumui tirti atlikta patvirtinamoji faktorinė analizė (angl. *CFA*). Rezultatai rodo, kad į aukštesnio lygio mąstymo įgūdžius orientuoti technologinių pedagoginių turinio žinių tyrimo instrumentai yra pagrįsti ir patikimi, penki validumo rodikliai priskirti kategorijai „tinka“, o vidutinis patikimumo koeficientas yra 0,85. Taigi, šis į aukštesnio lygio mąstymo įgūdžius orientuotas technologinių pedagoginių turinio žinių klausimynas (angl. *TPACK-HOTS*) gali būti naudojamas kaip parametras vertinant kalbų mokytojų kompetencijas.

Esminiai žodžiai: *aukštesnio lygio mąstymo įgūdžiai (angl. HOTS), kalbų mokytojai, apklausos rengimas, technologinės pedagoginės turinio žinios (angl. TPACK), validumas.*

Gauta 2022 06 04 / Received 04 06 2022
Priimta 2022 12 27 / Accepted 27 12 2022