Development and Validation of Technological Pedagogical and Content Knowledge – Action Research in Teaching Science Instrument

Arnel A. Lorenzana[1,2,*] and Lydia S. Roleda[2]

[1] Science, Technology, and Mathematics Department, Bicol University College of Education, Legazpi City, Philippines
[2] Department of Science Education, De La Salle University, Manila, Philippines
Email: arnel_a_lorenzana@dlsu.edu.ph (A.A.L.); Lydia.roleda@dlsu.edu.ph (L.S.R)
*Corresponding author

***Abstract*—This article describes the development and validation of Technological Pedagogical and Content Knowledge – Action Research in Teaching Science (TPACK-ARTS) Instruments. This instrument measures pre-service science teachers' perception of action research and TPACK. The development and validation underwent three phases: instrument design, content validation, and construct validation. The items in the instrument were designed based on literature, and nine experts examined the appropriateness of the items in the instrument. The data from the expert validation was analyzed using Lawshe's Content Validity Ration (CVR) method. The result suggested content validity. Exploratory factor analysis (N = 107) using principal component analysis and Varimax with Kaiser Normalization rotation generates three factors. Confirmatory factor analysis showed that the three-factor model satisfies all the eligibility criteria for fit indices, showing evidence of construct validity. Furthermore, reliability analysis using the composite reliability, average variance extracted, standardized factor loadings, Cronbach's alpha, and corrected item-total correlation all suggest the high reliability of the instrument.**

***Keywords*—Technological Pedagogical and Content Knowledge (TPACK), action research, instrument development, Content Validity Ratio (CVR), factor analysis**

I. Introduction

Developing Pre-service Teachers' (PSTs) competencies to accommodate the current demands in the basic educational setting is always a significant and vital issue. Fundamental to this concern is how to prepare PSTs to have a sound grasp of the subject's content, pedagogical competency, and the use of technology in teaching. Central to this is how to develop the interplay of these three forms of teachers' knowledge to teach the subject matter with technology successfully. Kafyulilo [1] pointed that effective teaching with technology requires that teachers understand the content they want to teach, the pedagogy concurrent with the subject's content, and the technology that can support students' learning in a specific context.

The program for Pre-service Science Teachers (PSSTs) in the Philippines is the Bachelor in Secondary Education major in science. It aims to produce PSSTs who can "(1) demonstrate a deep understanding of scientific concepts and principles, (2) apply scientific inquiry in teaching and learning, and (3) utilize effective science teaching and assessment methods" [2]. The program includes a variety of courses focusing on content knowledge, pedagogical knowledge, and technological knowledge to attain its goal. These courses include the Teaching Science and Action Research courses, which allow PSSTs to practice and investigate their Technological Pedagogical and Content (TPACK)-related competencies. Until now, however, little is known, especially in the Philippines, about how action research develops PSSTs' TPACK in Science-related skills and how they perceive action research in relation to TPACK. Examining PSSTs' perception of action research and how it relates to TPACK may help them better understand themselves as better science teachers in the future. In response, the present study developed and validated an instrument that can measure the level of pre-service science teachers' perception of their action research and TPACK-related skills.

II. Methods

The strength of an instrument is in the in-depth study of the instrument itself. The reliability and validity of an instrument are provided by a step-by-step approach [3]. The instrument development and validation in this study were undertaken in three phases. First, the Technological Pedagogical and Content Knowledge – Action Research in Teaching Science (TPACK-ARTS) instrument was developed through a review of related literature and interviews. Second, Content Validation by a panel of experts ensured the appropriateness of each instrument item to what it intended to measure. The last phase was Construct Validation and Reliability Analysis. This phase ensured that the instrument reflected the latent theoretical

Manuscript received May 4, 2023; revised July 27, 2023; accepted August 10, 2023; published February 2, 2024.

constructs it designed to measure, thus confirming internal reliability of the instrument.

A. Instrument Design

This stage consisted of three steps process that include determining the content domain, item generation, and instrument construction [4]. Shrotryia and Dhanda [5] mentioned that the content domain of the construct is determined by literature review, content analysis, and interviews. This step ensure that the attributes and agreed-upon definition of the construct are obtained and separated from others. A literature review was conducted to determine the content domain for TPACK-ARTS. The items generated were based on the reviewed literature and interviews. The items were initially identified along with three domains or constructs: Technological Pedagogical and Content Knowledge in Teaching Science (TPACK-TS), Action Research (AR), and Technological Pedagogical and Content Knowledge – Action Research in Teaching Science (TPACK-ARTS). A total of thirty-three items were initially constructed along with the three domains, TPACK-ST (9 items), AR (14 items), and TPACK-ARTS (10 items). The Likert method of ratings was chosen as an item response format with four options: SD – Strongly Disagree, D – Disagree, A – Agree, and SA – Strongly Agree.

B. Content Validation

This stage involves confirming the items by a specific number of experts to ensure the content validity of the assessment instrument. Content validity is vital to support the validity of an assessment tool such as questionnaires, especially for research purposes [6]. Content validity specifies if the items in the instrument sample the complete range of the attribute under study [7]. Experts were selected based on expert knowledge, specific training, or professional experience on the subject matter [5]. In this study, the initial draft of the instrument was given to a group of nine experts on science education, action research, educational foundation, and teacher education. These experts evaluated the selected items to determine whether it is "*Essential*", "*Useful but not essential*", or "*Not necessary*" in each of the three dimensions. Experts also provided suggestions for improving the statements and the instrument.

The content validity of each item was estimated using Lawshe's Content Validity Ratio (CVR) [8, 9]. The CVR is a proper statistical technique to determine the validity of individual instrument items, as rated by a panel of content experts [10]. Equation 1 shows how CVR is calculated. For a panel of nine experts, a CVR of 0.778 or higher could be considered evidence of content validity [5, 6, 9, 11, 12]. The Kappa coefficient was also estimated to ensure the agreement was not due to chance. Kappa coefficients higher than 0.74 have an excellent interrater agreement [4, 5, 13]. The content validity of the whole instrument was calculated using the Content Validity Index (CVI), which is simply the mean of the CVR values for all items meeting the CVR threshold of 0.78 and retained for the final instrument [10].

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \quad (1)$$

where: n_e is the number of experts identifying an item as essential, N is the total number of experts

C. Construct Validation

The instrument was given to 107 Pre-Service Science Teachers (PSSTs) from Teacher Education Institutions (TEI) in Bicol Region, Philippines. PPSTs comprised 36 males and 71 females from the first to third-year level. Exploratory Factor Analysis (EFA), Confirmatory Factor Analysis (CFA), and reliability analysis were conducted from the collected data from PSSTs. Exploratory factor analysis is a procedure that assists researchers in identifying or understanding "latent" constructs underlying variables of interest [14, 15]. According to Hair *et al.* [16], the statistical objective of EFA is to identify a set of latent constructs from several individual items. The purpose of EFA was to investigate the factors underlying the TPACK-ARTS in this study. Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity (BTS) were used to test if EFA can be performed on the selected sample. If the KMO value is more than 0.5 [17], more than 0.6 and higher [14, 18] are suggested for pushing forth with factor analysis. A statistically significant ($p < 0.001$) result of BTS suggests having a sufficient correlation to carry out factor analysis [14].

Exploratory factor analysis with the principal component method with Varimax with Kaiser Normalization as a rotation method was utilized to identify or examine the factor structure of the questionnaire. Scree test and Kaiser's (Eigenvalue) Criterion were used to determine the number of factors to be extracted. An eigenvalue greater than one is considered in determining the number of factors, while the factor load should be 0.40 or higher [19]. Parallel analysis was also done to determine the number of factors using the Parallel Analysis Engine available from https://analytics.gonzaga.edu/parallelengine/. Items with cross-loadings were omitted [3]. EFA was carried out using SPSS.

CFA was carried out using AMOS 21 on a second sample (n = 239) to examine the factor structure of TPACK-ARTS on findings from EFA. This method is used to see if the quality of the items developed can measure what it intends to measure [20, 21]. Furthermore, it is both a qualitative and statistical process that examines the reliability of the individual indicators (item reliability), construct reliability, quantitative measures of convergent and discriminant validity, and Goodness of Fit [16].

The data obtained from CFA were evaluated according to different fit indexes. According to Tabachnick and Fidell [22], the Comparative Fit Index (CFI) and Root Mean Square Error of Approximation (RMSEA) are perhaps the most frequently reported indices. They also mentioned that if the results of the fit indices are inconsistent, the model should be re-examined and consider reporting multiple indices if the inconsistency is not resolved.

Hair *et al.* [21] recommended that the eligibility criteria for fit indices be used to evaluate the model's validity in this study. They mentioned that multiple indexes should assess a model's goodness of fit. It should include the chi-squared and associated degree of freedom, one absolute fit index (i.e., Goodness of Fit Index (GFI), RMSEA, or Standard Root Mean Residual (SRMR), one incremental fit index (i.e., CFI or Tucker Lewis Index (TLI)), One goodness-of-fit index (GFI, CFI, TLI), one badness-of-fit index (RMSEA, SRMR). However, the chi-squared goodness of fit index test or resulting p-value is less meaningful, particularly as sample sizes become large or the number of observed variables becomes large. Therefore, in this study, only CFI, RMSEA, and SRMR were used to assess the goodness of fit indices of the instrument.

The Comparative Fit Index (CFI) is an increment fit index that is an improved version of the normed fit index. CFI supplements the chi-squared and df. The larger the CFI, the better the fit [22]. Root Mean Square Error Approximation (RMSEA) is a measure that attempts to correct for the tendency of the chi-squared goodness of fit test statistic to reject models with large samples or a large number of observed variables. Lastly, Standardized Root Mean Square (SRMR) helps compare fit across models. The smaller the value of RMSEA and SRMR, the better the fit.

Reliability analysis was also conducted to estimate the construct validity for each factor generated using Composite Reliability (CR), Average Variance Extracted (AVE), Cronbach's alpha, corrected item-total correlation, and standardized factor loadings. Reliability is a measure of the degree to which a set of indicators of a latent construct is internally consistent based on how highly interrelated the indicators are to each other [21].

According to Syahfitri *et al.* [20], the composite (construct) reliability method is a guide to seeing reliability analysis. CR can be conceptualized as how well the indicators represent the latent variable [23]. Composite reliability higher than 0.7 is needed in a more advanced validation phase [18]. On the other hand, AVE that is 0.5 or higher [18, 21] confirms convergent validity. Cronbach's alpha also measures internal consistency; a value of more than 0.7 is considered acceptable [24]. The corrected item-total correlation shows how each item correlates with the overall questionnaire. It expresses the coherence between an item and other items in a test [25]. A correlation less than $r = 0.30$ indicates that the item may not belong on a scale [26]. Another statistic that ensures reliability therefore reflecting construct validity, is the standardized factor loading. Individual weights should be greater than 0.5.

III. Findings and Discussion

A. *Content Validity*

A review of related literature informed the initial phase of developing and validating the TPACK-ARTS questionnaire (see [3, 27–42]). The questionnaire (33 items) was initially designed to have three factors or constructs. These are the Technological Pedagogical and Content Knowledge in Teaching Science (TPACK-ST), Action Research (AR) in an educational setting, and Technological Pedagogical and Content Knowledge – Action Research in Teaching Science (TPACK-ARTS) which described the relationship between AR and TPACK.

The findings in the CVR analysis showed that not all panels of nine experts agreed on all items in the three factors or constructs. An item considered valid should have at least eight out of nine experts agreeing that the item is essential. If an item did not reach this threshold, it was removed from the initially validated instrument [10]. As shown in Table I, six items (in *) did not reach the minimum standard; thus, were removed from the initial draft of the instrument. Only items with CVR equal to 0.78 or higher were retained. Items with a kappa coefficient less than 0.74 were removed to ensure interrater agreement. Incidentally, items with low kappa coefficients were also the six items with low CVR coefficients. Removing items this way followed Polit *et al.* [11] when they stated that considering adjusted kappa equal to or higher than 0.74 and CVR equal to or greater than 0.78 would be excellent.

The mean of all CVR per construct or factor was also calculated. Table I shows that TPACK, AR, and TPACK-ARTS got a CVI of 0.94, 0.84, and 0.88, respectively. The CVI obtained for the whole was equal to 0.88. These values are beyond the minimum standards, suggesting that the full scale has content validity.

B. *Exploratory Factor Analysis of the Instrument*

The purpose of exploratory factor analysis was to determine factors underlying TPACK-ARTS in this study. Principal Component Analysis was used to determine the data's suitability for factor analysis. The Kaiser-Meyer Olkin measure of sampling adequacy and Bartlett's Test of Sphericity have evaluated whether the sample was large enough. A KMO value of 0.6 and higher is suggested for pushing forth with the factor analysis [14, 22]. The KMO value of the initial analysis was 0.884, considered very good [22] and meritorious [21]. Bartlett's Test of Sphericity was significant; indicating that the sample and correlation matrix were appropriate for the factor analysis. The initial solution yielded four factors with an eigenvalue greater than 1. These factors altogether explained 65.358% of the variance result. It is suggested that the proportion of the total variance explained by retained factors should be at least 50% [15, 18]. Furthermore, other studies [3, 17, 28, 35] reported greater than 50% total variance explained.

TABLE I. CVR, CVI, AND KAPPA STATISTICS OF THE INSTRUMENT

Code		N_e	CVR	P_c	Kappa
Technological Pedagogical and Content (Science) Knowledge (TPACK)					
TPACK_1	I can design lessons that appropriately combine technology and teaching strategies on particular science concepts.	9	1.00	0.002	1.00
TPACK_2	I can teach lessons that appropriately integrate technologies and teaching strategies on particular science concepts.	9	1.00	0.002	1.00
TPACK_3	I can use technology to teach science effectively using various teaching strategies.	9	1.00	0.002	1.00
TPACK_4	I can select appropriate technologies that can simplify science concepts to enhance how I teach and what the students learn.	9	1.00	0.002	1.00
TPACK_5	I can design engaging learning activities that appropriately combine technologies and teaching strategies on particular science concepts.	9	1.00	0.002	1.00
TPACK_6	I can design individual or collaborative performance tasks that use technology to show, apply, and/or solve science-related problems.	9	1.00	0.002	1.00
TPACK_7	I know how to use appropriate technologies as a tool for sharing ideas and thinking together in the teaching-learning process in a science class.***	8	0.78	0.018	0.77
TPACK_8	I can evaluate the student's performance and understanding of science using appropriate technologies and assessment strategies.	8	0.78	0.018	0.77
TPACK_9	I can teach a particular science topic using different strategies and technologies.*	7	0.56	0.281	0.38
CVI	**0.94**				
Action Research (AR)					
AR_1	As a teacher, I know how to investigate my teaching practice.	9	1.00	0.002	1.00
AR_2	I reflect on my teaching practice and use it to plan a possible strategy to address concerns.**	8	0.78	0.018	0.77
AR_3	I can assume the roles of both the teacher and researcher at the same time.**	8	0.78	0.018	0.77
AR_4	I can identify issues (such as problems, obstacles, etc.) in my teaching practice.**	9	1.00	0.002	1.00
AR_5	I can make action plans based on the root causes of the problem of my teaching practice.	9	1.00	0.002	1.00
AR_6	I can carry out my action plans to address issues in my teaching practice.***	8	0.78	0.018	0.77
AR_7	I know how to gather relevant data to analyze the effectiveness of my plan.	8	0.78	0.018	0.77
AR_8	I am aware of the possible bias/es I may have in analyzing qualitative data.	8	0.78	0.018	0.77
AR_9	I know how to analyze data using statistical tools and make a generalization.	8	0.78	0.018	0.77
AR_10	I can systematically resolve teaching-related problems by applying the results of my action plan.	8	0.78	0.018	0.77
AR_11	I can set goals to improve my teaching skills.*	6	0.33	5.906	1.14
AR_12	I can carry out my plans to improve my teaching skills.*	7	0.56	0.281	0.38
AR_13	I know how to apply what I have learned to change my teaching practice.*	7	0.56	0.281	0.38
CVI	**0.84**				
Technological Pedagogical Content Knowledge - Action Research in Teaching Science (TPACK-ARTS)					
TPACK ARTS_1	I can improve my technological pedagogical content knowledge by doing action research.***	9	1.00	0.002	1.00
TPACK ARTS_2	I am aware that the goal of action research is to improve myself as a holistic science teacher.***	9	1.00	0.002	1.00
TPACK ARTS_3	I can improve the lessons by combining technologies, teaching strategies, and science concepts appropriately through action research.***	9	1.00	0.002	1.00
TPACK ARTS_4	I can improve my teaching strategies by reflecting on action research findings on different topics.	8	0.78	0.018	0.77
TPACK ARTS_5	I can design better assessment practices using technology through action research.	8	0.78	0.018	0.77
TPACK ARTS_6	I can improve my teaching practices through reflection on every part of the action research activity.	9	1.00	0.002	1.00
TPACK ARTS_7	I can apply the findings of action research to different scenarios of the teaching-learning process.	8	0.78	0.018	0.77
TPACK ARTS_8	I can better understand myself as a teacher through action research.	8	0.78	0.018	0.77
TPACK ARTS_9	I am open to changes as long as they can improve my teaching practice.***	8	0.78	0.018	0.77
TPACK ARTS_10	I feel I can better understand my students through action research.*	7	0.56	0.281	0.38
TPACK ARTS_11	I can improve students' engagement through action research.*	7	0.56	0.281	0.38
CVI	**0.88**				
SCALE CVR	**0.88**				

Note: Items with asterisks were removed during the validation procedure: *experts validation, **EFA, ***CFA analysis.

The scree plot in Fig. 1 shows that four factors were in sharp descent and then started to level off. It was shown that the largest eigenvalue was 11.670, followed by 2.763. The two remaining factors had an eigenvalue that was only more than one. This result required further analysis on whether to include other factors.

Exploratory Factor Analysis (EFA) was conducted using principal component analysis and Varimax with Kaiser Normalization rotation with four factors to be extracted. However, there was cross-loading among items. Therefore, an alternative approach was conducted using parallel analysis utilizing [43]. By comparison (see Table

II), three factors were more significant than the mean eigenvalue indicating three factors to be extracted. This outcome is understandable because theoretically, only three factors were suggested.

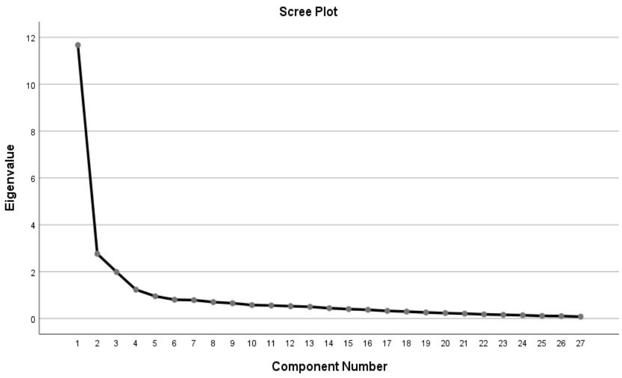

Fig. 1. Scree plot.

TABLE II. PARALLEL ANALYSIS

Component or Factor	Mean Eigenvalue (Patil *et al.* [43])	Derived Eigenvalue
1	2.071	10.665
2	1.886	2.655
3	1.753	1.875
4	1.651	1.018

TABLE III. FACTOR LOADINGS

	Component		
Code	1	2	3
TPACK_1		0.612	
TPACK_2		0.721	
TPACK_3		0.663	
TPACK_4		0.717	
TPACK_5		0.763	
TPACK_6		0.752	
TPACK_7		0.716	
TPACK_8		0.75	
AR_1			0.578
AR_5			0.771
AR_6			0.775
AR_7			0.753
AR_8			0.656
AR_9			0.594
AR_10			0.74
TPACK_ARTS_1	0.726		
TPACK_ARTS_2	0.76		
TPACK_ARTS_3	0.74		
TPACK_ARTS_4	0.89		
TPACK_ARTS_5	0.803		
TPACK_ARTS_6	0.718		
TPACK_ARTS_7	0.712		
TPACK_ARTS_8	0.767		
TPACK_ARTS_9	0.526		

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization
a. Rotation converged in 6 iterations.

Another run was performed, with only three factors to be extracted. Hair *et al.* [21] suggested that problematic and potential cross-loadings should be deleted. Three items with cross-loading were removed, and EFA was run for the third time. Table III shows three factors generated by the instrument. These factors explained 63.312% of the total variance. Factor loadings ranged from 0.526 to 0.89, which were considered significant, given that the sample size was 239. Factor loadings of 0.40 and higher were considered significant for interpretative purposes [21].

The distribution of the items among each component is grouped according to the proposed groupings. This result could be because a panel of experts already validated the items based on a review of related literature. The name of the factors was still the same as before.

C. Confirmatory Factor Analysis

Construct validity of the measurement instrument was estimated using confirmatory factor analysis. Construct validity is how a set of measured items accurately reflects the latent theoretical constructs designed to measure. Thus, construct validity deals with the accuracy of measurement. CFA aims to reveal the latent construct of the items in an instrument [44]. The three-factor model obtained through EFA was then analyzed through CFA using AMOS 21 application. In the first attempt, the fit indices of the model indicated the model did not fit well with CFI = 0.922 and RMSEA = 0.074. The modification indices were then checked. The residual of some items was correlated to each other; items with high standard residual covariances were deleted [21]. Several items were deleted in this process: TPACK_7, AR_1, AR_6, TPACK-ARTS_1-3, and 9. The path diagram of the three-factor model with the standardized coefficients between items and factors is shown in Fig. 2. Further assessment of the factor loading of each item (0.75–0.89) showed that no item is below the minimum (0.50).

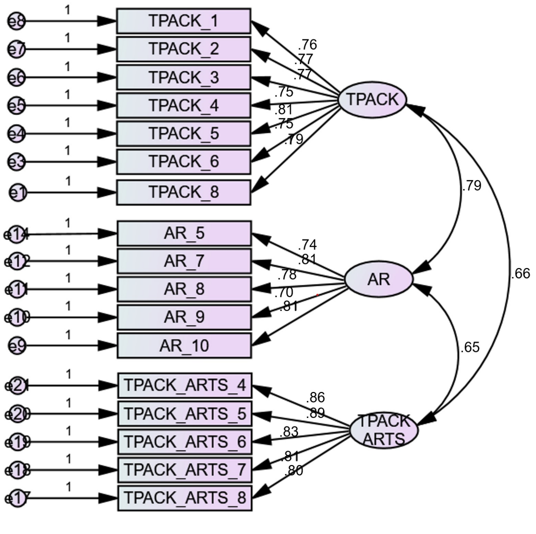

Fig. 2. Path diagram of the instrument.

The model fit measures were used to assess the model's overall goodness of fit (χ^2 and df, CFI, RMSEA, and SRMR), and all values are within their respective acceptance level [3, 15, 21, 22, 35, 45, 46]. The three-factor model (TPACK, AR, TPACK-ARTS) yielded a good fit (Table IV) for the data: χ^2 = 209.883 (df = 116, p = 0.000), Comparative Fit Index (CFI) = 0.97, Root Mean Square Residual (RMSEA) = 0.058, and Standardized Root Mean Square (SRMR) = 0.039.

TABLE IV. FIT INDICES OF THE FINAL CONSTRUCTION OF THE TPACK-ARTS INSTRUMENT

Goodness of Fit index	Recommended Value (Hair *et al.* [21])	TPACK-ARTS Instrument
X^2	Significant p-values, even with a good fit	Significant
CFI	0.97 or better	0.97
SRMR	0.08 or less with a CFI of 0.95 or higher)	0.039
RMSEA	Values < 0.08 with CFI of 0.97 or higher	0.058

Note: Recommended value for N < 250 and 12 < m < 30, where N is the number of observations and m is the number of observed variables.

D. *Reliability of the Instrument*

To ensure that the instrument is reliable, the internal consistency of the total scale and the three factors were estimated using Cronbach's Alpha (CA), composite rating (Construct Reliability) (CR), and Average Variance Extracted (AVE). Standardized factor LOADING (FL) and Corrected Item-Total Correlation (CITC) were also used to estimate how much each item correlated with the overall questionnaire. Table V describes the reliability estimates of the total scale, factors, and items. The Cronbach's alpha of the TPACK was 0.943, AR was 0.876, and TPACK-ARTS was 0.922. Also, it was found that Cronbach's alpha of the total scale was 0.943. The result was beyond the accepted value [18, 24, 47], showing that the scale had evidence of internal consistency. Composite Reliability for the three factors ranged from 0.879 to 0.922, which indicated that each construct in the three-factor model is reliable [21, 45]). The average variance extracted for each factor (0.536–0.771) was also within the normal range [21]. CITC ranged from 0.611 to 0.726. This meant the item belonged to the scale [17, 19]. All factor loading estimates were more significant than 0.70, suggesting convergence or internal consistency. All indices indicated that the instrument has internal consistency and construct validity.

TABLE V. MEAN, SD, AND RELIABILITY OF THE INSTRUMENT

Goodness of Fit index	Recommended Value (Hair *et al.* [21])	TPACK-ARTS Instrument
X^2	Significant p-values, even with a good fit	Significant
CFI	0.97 or better	0.97
SRMR	0.08 or less with a CFI of 0.95 or higher)	0.039
RMSEA	Values < 0.08 with CFI of 0.97 or higher	0.058

Note: Recommended value for N < 250 and 12 < m < 30, where N is the number of observations and m is the number of observed variables.

IV. CONCLUSION

This study developed and validated a measurement instrument for assessing pre-service science teachers' perception of Technological Pedagogical Content Knowledge – Action Research in Teaching Science (TPACK-ARTS). The instrument's initial form, which is based on literature and interviews, consisted of 33 items. These items were distributed on predetermined factors (TPACK, AR, and TPACK-ARTS). Initial validation of the items using Lawshe's content validity ratio (CVR = 0.78–1.00), content validity index (CVI = 0.88), and kappa statistic (Kappa = 0.77–1.00) showed that the instrument had acceptable content validity. Six items were deleted from the instrument at this stage. Exploratory Factor Analysis (EFA) was conducted using principal component analysis and Varimax with Kaiser normalization rotation to determine if the items load correctly with the predetermined factors. Three factors were generated, which explained 63.312% of the total variance. Three items were removed due to cross-loading, but other items remained in the predetermined factors: TPACK, AR, and TPACK-ARTS. After EFA, Confirmatory Factor Analysis (CFA) was performed, to test the model obtained from EFA. Results showed that the three-factor model was compatible with the data in terms of the fit indices: χ^2 = 209.883 (df = 116, p = 0.000), CFI = 0.97, RMSEA = 0.058, and SRMR = 0.039. Seven items were removed in this stage. The composite reliability (CR = 0.879 to 0.922), average variance extracted (AVE = 0.536–0.771), standardized factor loadings (FL > 0.70), Cronbach's alpha (α = 0.943 (scale)), α = 0.876–0.943 (factors), and corrected item-total correlation (CITC = 0.611–0.726) all suggest high-reliability of the instrument. The development and validation procedure discussed above showed that the Technological Pedagogical and Content Knowledge – Action Research in Teaching Science instruments have good psychometric properties. It can be used to measure pre-service science teachers' perception of action research, technological and pedagogical content knowledge, and its interaction with each other.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

A.A. Lorenzana developed and validated the instrument and wrote the paper. L.S. Roleda supervised and gave very insightful ideas throughout the conduct of the study. All authors had approved the final version.

ACKNOWLEDGMENT

The authors would like to thank the DOST SEI-CBPSME Scholarship for funding the publication of the paper.

REFERENCES

[1] A. C. Kafyulilo. (2018). Developing Technological Pedagogical Content Knowledge (TPACK) among pre-service science and mathematics teachers at DUCE: The role of activity-based learning. Papers in Education and Development. [Online]. Available: https://journals.udsm.ac.tz/index.php/ped/article/view/1461

[2] CHED Memorandum Order No. 75. (2017). [Online]. Available: https://ched.gov.ph/wp-content/uploads/2017/11/CMO-No.-75-s.-2017.pdf

[3] T. Kartal and Ö. Afacan, "Examining Turkish pre-service science teachers' technological pedagogical content knowledge (TPACK) based on demographic variables," *Journal of Turkish Science*

Education, vol. 14, no. 1, pp. 1–22, 2017. https://doi.org/10.12973/tused.10187a

[4] V. Zamanzadeh, M. Rassouli, A. Abbaszadeh, *et al.*, "Details of content validity and objectifying it in instrument development," *Nursing Practice Today*, vol. 1, no. 3, pp. 163–171, 2015.

[5] V. K. Shrotryia and U. Dhanda, "Content validity of assessment instrument for employee engagement," *SAGE Open*, vol. 9, no. 1, 2019. https://doi.org/10.1177/2158244018821751

[6] M. S. B. Yusoff, "ABC of content validation and content validity index calculation," *Education in Medicine Journal*, vol. 11, no. 2, pp. 49–54, 2019. https://doi.org/10.21315/eimj2019.11.2.6

[7] H. A. DeVon, M. E. Block, P. Moyle-Wright, *et al.*, "A psychometric toolbox for testing validity and reliability," *Journal of Nursing Scholarship: An Official Publication of Sigma Theta Tau International Honor Society of Nursing*, vol. 39, no. 2, pp. 155–164, 2007. https://doi.org/10.1111/j.1547-5069.2007.00161.x

[8] C. H. Lawshe, "A quantitative approach to content validity," *Personnel Psychology*, vol. 28, no. 4, pp. 563–575, 1957. https://doi.org/10.1111/J.1744-6570.1975.TB01393.X

[9] F. R. Wilson, W. Pan, and D. A. Schumsky, "Recalculation of the critical values for Lawshe's content validity ratio," *Measurement and Evaluation in Counseling and Development*, vol. 45, no. 3, pp. 197–210, 2012. https://doi.org/10.1177/0748175612440286

[10] G. E. Gilbert and S. Prion, "Making sense of methods and measurement: Lawshe's content validity index," *Clinical Simulation in Nursing*, vol. 12, no. 12, pp. 530–531, 2016. https://doi.org/10.1016/j.ecns.2016.08.002

[11] D. F. Polit, C. T. Beck, and S. V. Owen, "Is the CVI an acceptable indicator of content validity? Appraisal and recommendations," *Research in Nursing & Health*, vol. 30, no. 4, pp. 459–467, 2007. https://doi.org/10.1002/nur.20199

[12] C. Ayre and A. J. Scally, "Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation," *Measurement and Evaluation in Counseling and Development*, vol. 47, no. 1, pp. 79–86, 2014. https://doi.org/10.1177/0748175613513808

[13] D. F. Polit and C. T. Beck, "The content validity index: Are you sure you know what's being reported? Critique and recommendations," *Research in Nursing & Health*, vol. 29, no. 5, pp. 489–497, 2006. https://doi.org/10.1002/nur.20147

[14] D. J. Denis, *SPSS Data Analysis for Univariate, Bivariate, and Multivariate Statistics*, John Wiley & Sons, Inc, 2019.

[15] B. L. Bandalos and S. J. Finney, "Factor analysis exploratory and confirmatory," in *The Reviewer's Guide to Quantitative Methods in Social Sciences*, G. R. Hancock, L. M. Stapleton, and R. O. Mueller, Eds., Routledge, 2018, pp. 98–122. https://doi.org/10.4324/9781315755649

[16] J. F. Hair, M. C. Howard, and C. Nitzl, "Assessing measurement model quality in PLS-SEM using confirmatory composite analysis," *Journal of Business Research*, vol. 109, pp. 101–110, 2020. https://doi.org/10.1016/j.jbusres.2019.11.069

[17] M. S. B. Yusoff, A. F. A. Rahim, and M. J. Yaacob, "The development and validity of the Medical Student Stressor Questionnaire (MSSQ)," *ASEAN Journal of Psychiatry*, vol. 11, no. 1, 2010.

[18] N. Shrestha, "Factor analysis as a tool for survey analysis," *American Journal of Applied Mathematics and Statistics*, vol. 9, no. 1, pp. 4–11, 2021. https://doi.org/10.12691/ajams-9-1-2

[19] R. H. Aşılar, S. Gözüm, C. Çapık, *et al.*, "Reliability and validity of the Turkish form of the eight-item Morisky medication adherence scale in hypertensive patients," *Anadolu Kardiyoloji Dergisi: AKD = The Anatolian Journal of Cardiology*, vol. 14, no. 8, pp. 692–700, 2014. https://doi.org/10.5152/akd.2014.4982

[20] J. Syahfitri, H. Firman, S. Redjeki, *et al.*, "Development and validation of critical thinking disposition test in biology," *International Journal of Instruction*, vol. 12, no. 4, pp. 381–392, 2019.

[21] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate DATA analysis*, 8th ed., UK: Cengage, 2019.

[22] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th ed., NY: Pearson, 2019.

[23] G. R. Hancock and R. O. Mueller, "Rethinking construct reliability within latent variable systems," in *Structural Equation Modeling: Present and Future*, R. Cudeck, S. D. Toit, and D. Sörbom, Eds., Lincolnwood, IL: Scientific Software International, 2001, pp. 195–216.

[24] Y. X. Wang, X. M. Gu, and S. F. Liu, "The investigation and analysis of pre-service teachers TOWARD TPACK competencies," *Open Journal of Social Sciences*, vol. 8, pp. 327–339, 2020. https://doi.org/10.4236/jss.2020.812027

[25] R. Tapsir, N. A. N. Pa, and S. N. A. B. S. Zamri, "Reliability and validity of the instrument measuring values in mathematics classrooms," *Malaysian Online Journal of Educational Sciences*, vol. 6, no. 2, pp. 37–47, 2018.

[26] E. A. O. Zijlmans, J. Tijmstra, L. A. V. D. Ark, and K. Sijtsma. (2019). Item-score reliability as a selection tool in test construction. *Frontiers in Psychology.* [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02298/full

[27] D. A. Schmidt, E. Baran, A. D. Thompson, *et al.*, "Technological Pedagogical Content Knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers," *Journal of Research on Technology in Education*, vol. 42, no. 2, pp. 123–149, 2009. https://doi.org/10.1080/15391523.2009.10782544

[28] Z. Aksu, M. Metin, and A. C. Konyalıoğlu, "Development of the pedagogical content knowledge scale for pre-service teachers: The validity and reliability study," *Mediterranean Journal of Social Sciences*, vol. 5, no. 20, 1365, 2014.

[29] T. Valtonen, E. T. Sointu, K. Mäkitalo-Siegl, *et al.* (2015). Developing a TPACK measurement instrument for 21st century pre-service teachers. [Online]. Available: https://journals.oslomet.no/index.php/seminar/article/view/2353

[30] M. Ulvik and H. Riese, "Action research in pre-service teacher education – A never-ending story promoting professional development," *Professional Development in Education*, vol. 42, no. 3, pp. 441–457, 2016. https://doi.org/10.1080/19415257.2014.1003089

[31] S. Kennedy-Clark, K. Eddles-Hirsch, T. Francis, *et al.*, "Developing pre-service teacher professional capabilities through action research," *Australian Journal of Teacher Education*, vol. 43, no. 9, pp. 39–58, 2018. https://search.informit.org/doi/10.3316/aeipt.221584

[32] I. Laudonia, R. Mamlok-Naaman, S. Abels, *et al.*, "Action research in science education – An analytical review of the literature," *Educational Action Research*, vol. 26, no. 3, pp. 480–495, 2018. https://doi.org/10.1080/09650792.2017.1358198

[33] M. S. Prudente and S. E. Aguja, "Principles, attitudes and processes in doing action research: Perceptions of teachers from the Philippines," *Advanced Science Letters*, vol. 24, no. 11, pp. 8007–8010, 2018. https://doi.org/10.1166/asl.2018.12478

[34] K. Açıkgül and R. Aslaner, "Investigation relations between the technological pedagogical content knowledge efficacy levels and self-efficacy perception levels of pre-service mathematics teachers," *Cukurova University Faculty of Education Journal*, vol. 48, no. 1, pp. 1–31, 2019.

[35] M. Irmak and Ö. Y. Tüzün, "Investigating pre-service science teachers' perceived Technological Pedagogical Content Knowledge (TPACK) regarding genetics," *Research in Science and Technological Education*, vol. 37, no. 2, pp. 127–146, 2019. https://doi.org/10.1080/02635143.2018.1466778

[36] M. P. E. Morales, J. C. Anito, R. A. Avilla, *et al.*, *TPACK in Philippine STEAM Education*, Manila, Philippines: Philippine Normal University, 2019.

[37] J. Maksimović and J. Osmanović, "Perspective of cognitive thinking and reflective teaching practice," *International Journal of Cognitive Research in Science, Engineering and Education*, vol. 7, no. 2, pp. 1–10, 2019. https://doi.org/10.5937/IJCRSEE1902001M

[38] D. Saribas and Z. G. Akdemir, "Action research on pre-service elementary teachers' understandings of the scientific method and the use of evidence in a science and technology teaching course," *Research in Science & Technological Education*, 2020. https://doi.org/10.1080/02635143.2020.1814233

[39] M. Schmid, E. Brianza, and D. Petko, "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 and Education*, vol. 157, 103967, 2020. https://doi.org/10.1016/j.compedu.2020.103967

[40] Z. K. Ünlü and İ. Dökme, "The effect of technology-supported inquiry-based learning in science education: Action research," *Journal of Education in Science Environment and Health*, vol. 6, no. 2, pp. 120–133, 2020. https://doi.org/10.21891/jeseh.632375

[41] P. Y. Sarpkaya and B. Altun, "Professional qualification scale for pre-service teachers," *E-International Journal of Educational Research*, vol. 12, no. 1, pp. 104–123, 2021.

[42] M. Schmid, E. Brianza, and D. Petko. (2021). Self-reported Technological Pedagogical Content Knowledge (TPACK) of pre-service teachers in relation to digital technology use in lesson plans. *Computers in Human Behavior*, Elsevier. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563220303332

[43] V. H. Patil, S. N. Singh, S. Mishra, and D. T. Donavan. (2017). Parallel analysis engine to aid in determining number of factors to retain using R [computer software]. [Online]. Available: https://analytics.gonzaga.edu/parallelengine/

[44] T. A. Brown, *Confirmatory Factor Analysis for Applied Research*, New York: Guilford Publications, 2015.

[45] Y. C. Huang and S. H. Lin, "Assessment of charisma as a factor in effective teaching," *Educational Technology and Society*, vol. 17, no. 2, pp. 284–295, 2014.

[46] R. Gomez, V. Stavropoulos, and M. D. Griffiths, "Confirmatory factor analysis and exploratory structural equation modelling of the factor structure of the depression anxiety and stress scales-21," *PLOS ONE*, vol. 15, no. 6, e0233998, 2020. https://doi.org/10.1371/journal.pone.0233998

[47] K. S. Taber, "The use of Cronbach's alpha when developing and reporting research instruments in science education," *Research in Science Education (Australasian Science Education Research Association)*, vol. 48, no. 6, pp. 1273–1296, 2017. https://doi.org/10.1007/s11165-016-9602-2