Jurnal Ilmiah Sekolah Dasar
Volume 8, Number 3, 2024 pp. 601-611
P-ISSN: 2579-3276 E-ISSN: 2549-6174
Open Access: https://doi.org/10.23887/jisd.v8i3.81105

Enhancing Prospective Elementary School Teachers' 4C Skills: Instrument Development in the Context of Materials and Changes Lecture

Yusup Maulana[1], Wahyu Sopandi[2*], Rahmat Iqbal[3], Maryos Ipaubla[4]

[1,2] Graduate School, Elementary Education Study Program, Universitas Pendidikan Indonesia, Bandung, Indonesia
[3] Faculty of Teacher Training and Education, Elementary School Teacher Education Program, Universitas Syiah Kuala, Aceh, Indonesia
[4] Faculty of Teacher Training and Education, Elementary School Teacher Education Program, Universitas Pattimura, Ambon, Indonesia

ARTICLE INFO

Article history:
Received January 24, 2024
Accepted August 12, 2024
Available online August 25, 2024

Kata Kunci:
Keterampilan 4C, Pengembangan Instrumen, Reliabilitas, SmartPLS, *Validitas*

Keywords:
4C Skills, Instrument Development, Reliability, SmartPLS, Validity

ABSTRAK

Penelitian ini didorong oleh kebutuhan mendesak untuk memastikan bahwa calon guru sekolah dasar memiliki keterampilan 4C (Komunikasi, Kolaborasi, Berpikir Kritis, dan Kreativitas) yang memadai dalam menghadapi tantangan pendidikan abad ke-21, khususnya dalam konteks perkuliahan materi dan perubahan. Tujuan dari penelitian ini adalah untuk mengembangkan dan memvalidasi Instrumen keterampilan 4C Calon Guru Sekolah Dasar dalam perkuliahan topik materi dan perubahannya. Pendekatan yang digunakan adalah kuantitatif dengan desain survei yang melibatkan 400 mahasiswa pendidikan guru sekolah dasar di Indonesia. Pemilihan sampel didasarkan pada teknik pengambilan sampel acak sederhana. Temuan dari analisis PLS-SEM dari model pengukuran menunjukkan bahwa nilai uji validitas konvergen melalui factor loading untuk keterampilan komunikasi (CMC), kolaborasi (CLB), berpikir kritis (CTC), kreatifitas (CTV), dan tugas dosen (C) menunjukkan nilai lebih besar dari 0.6 dan dan nilai Average Variance Extracted (AVE) menunjukkan nilai lebih besar dari 0.5. Nilai uji Reliabilitas melalui nilai composite Reliability (CR) untuk setiap variabel menunjukkan nilai lebih besar dari 0.7 dan nilai Cronbach's Alpha menujukkan nilai lebih besar dari 0.6. Berdasarkan nilai tersebut instrumen dinyatakan valid dan reliabel sehingga instrumen dapat digunakan.

ABSTRACT

This research is driven by the urgent need to ensure that prospective elementary school teachers possess adequate 4C skills (Communication, Collaboration, Critical Thinking, and Creativity) to meet the challenges of 21st-century education, particularly in the context of materials and changes lectures. The purpose of this study was to develop and validate a 4C skills instrument for prospective elementary school teachers in lectures on the topic of matter and its changes. The approach used was quantitative with a survey design involving 400 elementary school teacher education students in Indonesia. The sample selection was based on simple random sampling technique. Findings from the PLS-SEM analysis of the measurement model showed that the convergent validity test value through factor loading for communication skills (CMC), collaboration (CLB), critical thinking (CTC), creativity (CTV), and lecturer assignments (C) showed values greater than 0.6 and and the Average Variance Extracted (AVE) value showed values greater than 0.5. The Reliability test value through the composite Reliability (CR) value for each variable shows a value greater than 0.7 and the Cronbach's Alpha value shows a value greater than 0.6. Based on this value, the instrument is declared valid and reliable so that the instrument can be used.

1. INTRODUCTION

In the era of globalization and rapid technological advancement, 21st century skills are becoming increasingly important for educators to master, especially prospective elementary school teachers. Among the various skills, the 4C skills, which include communication, collaboration, critical thinking, and creativity, are the main focus because they are considered essential in preparing students to face future challenges (Alismail, 2023; Widodo & Wardani, 2020). Given the importance of these skills, there is a need for valid and reliable instruments to measure the extent to which prospective teachers have mastered the 4C skills

Corresponding author
*E-mail addresses: wsopandi@upi.edu (Wahyu Sopandi)

(Fitriati et al., 2023; Maulana et al., 2021). A good measurement instrument must be reliable in various situations and contextualizations, so the development and validation of 4C skills instruments for prospective elementary school teachers is crucial (Nahar et al., 2022; Trisnawati & Sari, 2019). Given the importance of these skills, there is a need for valid and reliable instruments to measure the extent to which prospective teachers have mastered the 4C skills. A well-designed measurement instrument must be dependable in various situations and contexts, making the development and validation of a 4C skills assessment instrument for prospective elementary school teachers highly critical.

The 4C skills (Critical Thinking, Creativity, Collaboration, Communication) are particularly important for prospective elementary school teachers, who are students in Primary School Teacher Education programs (Maulana & Sopandi, 2022; Zubaidah, 2018). Educators have a crucial role in assessing students' 4C skills to determine how well these skills have developed and to identify areas that need improvement. This task includes planning and implementing comprehensive and continuous evaluations, using various methods such as observations, project assessments, portfolios, and performance tests. Furthermore, educators must provide constructive feedback to help students identify their strengths and weaknesses in 4C skills and guide them toward improvement strategies (Bialik et al., 2016; Erdoğan, 2019; Nahar et al., 2022). Recent research suggests that assessing 4C skills is not only important for individual student development but also to ensure they are prepared to meet the ever-changing challenges and demands of the education field. For instance, a study by other study indicates that the use of authentic assessments can significantly enhance students' critical thinking and collaborative skills (Griffin & Care, 2015; Koh et al., 2019). Meanwhile, research by other study emphasizes the importance of continuous assessment in monitoring the development of creativity and communication among students.

Additionally, educators need to integrate technology into the evaluation process to more effectively measure 4C skills. Technology can provide richer data and allow for more in-depth analysis of students' skill development. Integrating technology into education helps create a learning environment that supports the development of 21st-century skills, including the 4C skills (Laar et al., 2020; Nahar et al., 2022). Thus, the task of assessing the 4C skills of Primary School Teacher Education students is a complex and multidimensional process, requiring a holistic evaluation approach, the use of technology, and the provision of continuous and constructive feedback (Chen & Zhang, 2019; Khahro & Javed, 2022; Molla et al., 2023; Sunardi & Doringin, 2020). This research aims to address this gap by employing the SmartPLS (Partial Least Squares Structural Equation Modelling) approach in developing and validating the intended instrument. The instrument developed should be able to measure communication, collaboration, critical thinking, and creativity skills accurately and consistently.

PLS-SEM is a suitable method to use in instrument development due to its ability to handle complex models and relatively small sample sizes (Hair et al., 2017; Sarstedt et al., 2020). This method allows researchers to evaluate the relationship between indicators and measured constructs more accurately. In the context of education, the use of PLS-SEM can provide a clearer picture of the validity and reliability of the instruments used (Danks et al., 2020; Zhang et al., 2023). In this study, 400 elementary school teacher education students in Indonesia were selected as samples using simple random sampling techniques. The selection of this representative sample aims to ensure that the research results can be generalized to a wider population. The instrument developed should be able to measure communication, collaboration, critical thinking, and creativity skills accurately and consistently.

The convergent validity of this instrument is tested through factor loading analysis, where the resulting value must be greater than 0.6 to be considered valid. In addition, an Average Variance Extracted (AVE) value greater than 0.5 indicates that the measured construct is able to explain most of the variability of existing indicators. These results indicate that the instrument developed has good validity. Instrument reliability was also tested using the composite reliability (CR) and Cronbach's Alpha values. CR values greater than 0.7 and Cronbach's Alpha values greater than 0.6 indicate that this instrument has good internal consistency. This means that the instrument can be used consistently in various situations to measure 4C skills (Hidayatullah et al., 2021; Rodrigues et al., 2024). The development of a valid and reliable instrument is expected to be a useful tool for educational institutions in evaluating and improving their learning programs. By using this instrument, educational institutions can identify areas that need improvement and ensure that prospective teachers have the necessary skills to support effective learning processes (Basaria et al., 2021; Bray et al., 2023). In addition, the results of this study also make an important contribution to the field of educational research, particularly in the context of measuring 21st century skills. With a valid and reliable instrument, other researchers can use it for further studies and develop a deeper understanding of 4C skills development in teacher education. Overall, this research is expected to make a significant contribution to improving the quality of teacher education in Indonesia. Using the SmartPLS analysis approach, this research not only develops a solid measurement instrument, but also provides a solid foundation for the development of more data-driven education policies. Through rigorous validation and in-depth analysis, this instrument can be a useful guide for policy makers and education practitioners in their efforts to improve 4C skills among prospective primary school teachers. Based on this background, this study aims to develop and validate a 4C skills instrument for prospective elementary school teachers in lectures on the topic of matter and its changes. The novelty of this research highlights the development of 4C skills (Critical Thinking, Communication, Collaboration, and Creativity) specifically for prospective elementary school teachers. This focus has rarely been applied in previous, more general research on 4C training at the general student level or other occupational fields.

2. METHOD

This research uses a quantitative design with a survey research method. Survey research is research that uses questionnaires or interviews to collect data from a representative sample of a particular population to describe the characteristics, attitudes, or behavior of that population (Creswell, 2012). The sampling technique was random sampling with the aim of primary school teacher education students in Indonesia. The distribution of respondents came from universities that have elementary school teacher education study programs including universities from Aceh, Lampung, Bandung, Cibiru, Malang, Trenggalek, Bali, Ambon, Kalimantan, and Papua. A total of 400 respondents have filled out the questionnaire and the data has been reduced and can be used for analysis. The 4C skills research instrument was developed by compiling indicators of each construct, namely communication skills, collaboration skills, critical thinking skills, creativity, and tasks given by lecturers during lectures on the topic of matter and its changes. Each construct has 3 indicators, each indicator developed one statement with a 4-scale answer choice with the interpretation of "often", "sometimes", "rarely", and "never". The statement was coded according to the construct and given a number 1 to 3 for each indicator. Table 1 shows a description of the indicators of each construct developed.

Table 1. Number of Items in the Instrument

Construct	Indicator	Total	Code
Communication	Communicability: How comfortable you are in communicating with the lecturer and classmates about the concepts of matter and its changes.	1	CMC1
	Expression of Ideas in Discussion: How to express ideas or thoughts in class or study group discussions on the topics of matter and its changes.	1	CMC2
	Ability to Explain Concepts to Others: Skill in explaining concepts of matter and its changes to others who do not have the same knowledge.	1	CMC3
Collaboration	Participation in Group Discussions or Projects: Frequency of participation in group discussions or projects that involve problem solving in the context of matter and its changes.	1	CLB1
	Collaboration Ability: Feelings about the ability to effectively collaborate with team members to achieve common goals related to the topics of matter and change.	1	CLB2
	Handling Conflict or Dissent: How to handle conflict or disagreement in the context of group work in a course on matter and change.	1	CLB3
Critical Thinking	Use of Logical Reasoning: The level of use of logical reasoning in solving problems of matter and its changes in lectures.	1	CTC1
	Ability to Identify and Evaluate Arguments: The ability to identify and evaluate arguments that arise in discussions or lecture materials.	1	CTC2
	Confidence in Dealing with Complex Problems: The level of confidence in the ability to deal with complex problems in the context of matter and its changes.	1	CTC3
Creativity	Use of Creative Thinking: The frequency of using creative thinking to find alternative solutions to problems related to the concepts of matter and its changes.	1	CTV1
	Exploration of New Approaches or Innovative Ideas: The frequency of trying new approaches or innovative ideas in studying the course material matter and its changes.	1	CTV2

Construct	Indicator	Total	Code
	Ability to Generate Unusual Ideas or Solutions: How often you feel able to generate unusual ideas or solutions in the context of experiments or tasks related to matter and its changes.	1	CTV3
Coursework on Materials and Changes	The task of expressing your ideas or thoughts in class or study group discussions on the topics of matter and its changes (Part of collaboration and communication skills).	1	C1
	Tasks train logical reasoning, problem solving and evaluating arguments that arise in discussions (Part of critical thinking skills).	1	C2
	The task of finding ideas for investigations, problems, or creating works that are appropriate to the material that has been learned and related to everyday life (Part of creative thinking).	1	C3

*Table Description: Construct: Concepts or characteristics of the skills measured in the study; Indicator: Specific markers used to measure the construct; Total: Total number of indicators that have been measured for each construct; Code: A unique code used to identify each construct and indicator briefly and clearly.

Data were analyzed using Partial Least Squares Structural Equation Modeling (PLS-SEM) including Confirmatory Factor Analysis (CFA). In addition, SmartPLS 3.0 software was also used in this study. Although measurement and structural models can be tested simultaneously through a one-stage approach, SEM analysis is usually conducted using a two-stage approach. This approach involves the measurement model validation process using factor analysis (CFA) and the Structural Model Verification process (Henseler et al., 2015; Mardhotillah & Yulhendri, 2022). Measurement validation of the model involves four views: i. Identifying Goodness of Fit (GOF) values for each latent variable; ii. Identifying GOF values for the overall model measurements; iii. Correcting the GOF value of the Measurement model; iv. Determining the validity and reliability of the Measurement model. In the second stage of SEM analysis, testing of the Structural Model needs to be done by evaluating GOF and testing the proposed hypotheses. The data analysis carried out includes: Convergent Validity Test, Discriminant Validity Test, Reliability Test, Structural Model Test (R2 Square and Predictive Relevance), Significance Test (Path Coefficient or path analysis) (Gold et al., 2001; Sarstedt et al., 2020; Vinzi et al., 2010). Table 2 shows the validity and reliability values of the measurement model.

Table 2. Provisions of Validity and Reliability Values of the Measurement model

Parameters	Description	Acceptance Criteria
Convergent Validity Test	Factor Loading value for each indicator	Valid if the Factor loading value for each indicator is above 0.6 (>0.6)
	AVE (Average Variance Extracted) value for each variable	Valid if the AVE value for each variable is above 0.5 (>0.5)
Discriminant Validity Test	Cross Loading	It is passed if the value of the indicator is higher than other indicators
	Fornell Larcker	It is passed if the value of the indicator is higher than other indicators
	HTMT Value	Declared passed if the value is below 0.9 (<0.9)
Reliability Test	Composite Reliability Value	Reliable if the Composite Reliability value for each variable is above 0.7 (>0.7)
	Cronbach's Alpha Value	Reliable if the Cronbach's Alpha value for each variable is above 0.6 (>0.6)

3. RESULT AND DISCUSSION

Result

It is a challenge to select a suitable statistical model for data analysis. In this study, PLS-SEM analysis was used to obtain the research results for the fourth part. The choice of this analysis is based on the criterion that the data of this study is normally distributed. In addition, this study also uses a reflective-formative model and involves moderator variables, namely Continues. Table 3 shows the Cronbach's Alpha and CR values for each construct in this study which indicates that the 4C skills of prospective elementary school teachers can be explained using the developed instrument.

Table 3. Reliability Test

Construct	Cronbach's Alpha (>0.6)	Composite Reliability (>0.7)
Communication Skill (CMC)	0.658	0.815
Collaboration Skill (CLB)	0.660	0.815
Critical Thinking Skill (CTC)	0.636	0.805
Creativity (CTV)	0.690	0.829
Coursework on Materials and changes (C)	0.654	0.811

Table 3 presents the results of the reliability test analysis using Smart PLS for various constructs under examination. Five constructs were analyzed: Communication Skills (CMC), Collaboration Skills (CLB), Critical Thinking Skills (CTC), Creativity (CTV), and the Course on Materials and Changes (C). The Cronbach's Alpha values for all constructs are above 0.6, indicating that all constructs are reliable based on this criterion. Specifically, the Cronbach's Alpha values are 0.658 for Communication Skills, 0.660 for Collaboration Skills, 0.636 for Critical Thinking Skills, 0.690 for Creativity, and 0.654 for the Course on Materials and Changes. Additionally, the Composite Reliability values for all constructs exceed 0.7, further indicating that all constructs are reliable based on this criterion. Specifically, the Composite Reliability values are 0.815 for Communication Skills, 0.815 for Collaboration Skills, 0.805 for Critical Thinking Skills, 0.829 for Creativity, and 0.811 for the Course on Materials and Changes. Based on these results, it can be concluded that all the constructs tested exhibit good reliability, as indicated by both Cronbach's Alpha and Composite Reliability values. Predictive Relevance test (Q^2) is a test carried out in showing how well the resulting observation value is. If the Q^2 value < 0 then the variables and data have not been able to predict the model well, whereas if the Q^2 value > 0 indicates the variables and data can predict the model well (Vinzi et al., 2010). Table 4 shows the value of Predictive Relevance (Q^2) is more than 0 so that variables and data can predict the model well, which is similar or in accordance with the reality that occurs in the field.

Table 4. Predictive Relevance Value (Q^2)

Construct	Q^2 (=1-SSE/SSO)
Collaboration Skill (CLB)	0.111
Communication Skill (CMC)	0.084
Creativity (CTV)	0.163
Critical Thinking Skill (CTC)	0.050

Table 4 presents the results of the Predictive Relevance (Q^2) test for four constructs examined: Collaboration Skills (CLB), Communication Skills (CMC), Creativity (CTV), and Critical Thinking Skills (CTC). This Predictive Relevance test was conducted to assess how well the model can predict the observed values. The Q^2 values for each construct show positive results: Collaboration Skills have a Q^2 value of 0.111, Communication Skills 0.084, Creativity 0.163, and Critical Thinking Skills 0.050. All these Q^2 values are greater than 0, indicating that the variables and data can predict the model well. This means that the model used has good predictive capability, and the observed outcomes align with real-world situations. Therefore, it can be concluded that the model employed in this study has good predictive relevance for all the constructs tested. This analysis also uses convergent validity to assess how effective an item is in measuring similar constructs in a study Table 5 shows the results of the Outer Loading, CR, and AVE values. In this study, some items were excluded based on the requirements of the AVE value for each construct which must be more than 0.5 and the CR value which is more than 0.7. At the end of the analysis after several items were checked and eliminated that were not appropriate, all items in the table met the standard of meeting the minimum Outer Loading and AVE requirements.

Table 5. Convergent Validity Test

Construct	Item	Outer Loading (>0.6)	AVE (>0.5)
Communication (CMC)	CMC1	0.747	0.594
	CMC2	0.797	
	CMC3	0.768	
Collaboration (CLB)	CLB1	0.781	0.596
	CLB2	0.823	
	CLB3	0.708	
Critical Thinking (CTC)	CTC1	0.711	0.579
	CTC2	0.780	
	CTC3	0.790	
Creativity (CTV)	CTV1	0.785	0.618

Construct	Item	Outer Loading (>0.6)	AVE (>0.5)
Coursework on Materials and Changes (C)	CTV2	0.795	0.589
	CTV3	0.777	
	C1	0.817	
	C2	0.696	
	C3	0.785	

Table 5 presents the results of the convergent validity test, which aims to assess how effectively an item measures similar constructs within a study. The evaluation of convergent validity was conducted through three tests: Outer Loading Assessment, Composite Reliability (CR), and Average Variance Extracted (AVE). According to this analysis, the Outer Loading values should exceed 0.6, and the AVE values should be greater than 0.5 to meet the standards for convergent validity. The analysis results indicate that all items in the table have Outer Loading values greater than 0.6, with some values exceeding 0.7. Specifically, items such as CMC1, CMC2, CMC3 for the Communication Skills (CMC) construct; CLB1, CLB2, CLB3 for the Collaboration Skills (CLB) construct; CTC1, CTC2, CTC3 for the Critical Thinking Skills (CTC) construct; CTV1, CTV2, CTV3 for the Creativity (CTV) construct; and C1, C2, C3 for the Course on Materials and Changes (C) construct all demonstrate strong Outer Loading values. Additionally, the AVE values for each construct are also above 0.5, with Communication Skills at 0.594, Collaboration Skills at 0.596, Critical Thinking Skills at 0.579, Creativity at 0.618, and the Course on Materials and Changes at 0.589. Therefore, it can be concluded that all constructs in this study meet the requirements for convergent validity, based on both Outer Loading and AVE values, indicating that these items effectively measure the intended constructs.

Discriminant validity analysis is performed to assess the extent to which the tested constructs differ from other constructs. This analysis determines how much correlation between one construct and another and how many items can represent one particular construct (Hair et al., 2016). This study uses three methods to measure discriminant validity: 1) Cross Loading; 2) Fornell & Larcker Criteria; and 3) Heterotrait-Monotrait Ratio (HTMT).

The loading value for a particular construct must be greater than the loading value on other constructs (Hair et al., 2016). This indicates the subjective freedom of each predictor to the latent variable. This criterion helps reduce multicollinearity among latent variables by ensuring that the Average Variance Extracted (AVE) value of the latent variable is higher than all other variables (Chin, 1998; Fornell & Larcker, 1981; Vinzi et al., 2010). If the loading value for another construct is higher than the loading value for that construct, this indicates a discriminant validity problem (Sarstedt et al., 2020). Table 6 shows the results of the study indicate that the cross loading value for an item in a construct is greater than the loading value for other constructs. This finding provides validity evidence for the construct measurement model.

Table 6. Cross Loading

Code	Communication Skill (CMC)	Collaboration Skill (CLB)	Critical Thinking Skill (CTC)	Creativity (CTV)	Coursework on Materials and changes (C)
CMC1	0.747	0.406	0.391	0.370	0.289
CMC2	0.797	0.430	0.420	0.377	0.288
CMC3	0.768	0.411	0.532	0.433	0.308
CLB1	0.420	0.781	0.287	0.381	0.345
CLB2	0.491	0.823	0.357	0.359	0.367
CLB3	0.326	0.708	0.328	0.247	0.300
CTC 1	0.355	0.270	0.711	0.400	0.228
CTC2	0.431	0.274	0.780	0.408	0.218
CTC3	0.537	0.404	0.790	0.455	0.246
CTV1	0.375	0.383	0.492	0.785	0.409
CTV2	0.364	0.312	0.372	0.795	0.400
CTV3	0.466	0.319	0.441	0.777	0.409
C1	0.374	0.442	0.224	0.399	0.817
C2	0.209	0.218	0.223	0.350	0.696
C3	0.276	0.318	0.257	0.439	0.785

Table 6 shows the results of the study which indicate that the cross loading value for each item in a construct is greater than the loading value for other constructs. For example, item CMC1 has a loading value of 0.747 on the Communication Ability (CMC) construct, which is greater than its loading value on other constructs (0.406 for CLB, 0.391 for CTC, 0.370 for CTV, and 0.289 for C). Similarly, item CLB1 has a loading value of 0.781 on the Collaboration Ability (CLB) construct, which is greater than its loading value on other constructs (0.420 for CMC, 0.287 for CTC, 0.381 for CTV, and 0.345 for C). The same is true for other items, such as CTC1, CTV1, and C1, each of which has the highest loading value on the construct it is supposed to measure compared to other constructs. These findings provide evidence of strong discriminant validity for the construct measurement model, suggesting that the items used in this study are consistently more representative of the intended construct compared to other constructs. This suggests that the constructs tested in this study are truly distinct from one another, which is important for ensuring accurate and reliable interpretation of the research results.

The Fornell & Larcker criterion is an analysis that compares the square root value of AVE with the construct correlation value which shows the highest value in each column or row compared to the highest correlation value of other constructs (Fornell & Larcker, 1981). This approach is based on the opinion that latent variables should be better at explaining item variation than variation from other latent variables. After eliminating some indicators that do not meet the outer loading condition. Table 7 shows a higher squared AVE value compared to the correlation value for each other construct. The analysis results show that the Fornell-Larcker criterion in testing discriminant validity provides an answer to the research question regarding the validity of the construct measurement model.

Table 7. Fornell Lacker Criterion Value

Variable	Collaboration Skill (CLB)	Communication Skill (CMC)	Coursework on Materials and changes (C)	Creativity (CTV)	Critical Thinking Skill (CTC)
Collaboration Skill (CLB)	0.772				
Communication Skill (CMC)	0.539	0.771			
Coursework on Materials and changes (C)	0.438	0.383	0.768		
Creativity (CTV)	0.430	0.512	0.517	0.786	
Critical Thinking Skill (CTC)	0.419	0.583	0.304	0.554	0.761

Table 7 presents the results of the discriminant validity analysis using the Fornell-Larcker criterion. Each value on the bolded diagonal is the square root of the AVE (Average Variance Extracted) of each latent variable. This value is compared to the correlation value between other variables in the model. The Collaboration Skill (CLB) variable has an AVE value of 0.772, which is greater than its correlation with other variables, with the highest correlation of 0.539 with Communication Skill (CMC). Communication Skill (CMC) shows an AVE value of 0.771, which is also greater than the highest correlation of 0.583 with Critical Thinking Skill (CTC). Coursework on Materials and changes (C) has an AVE value of 0.768, which is higher than its highest correlation of 0.438 with Collaboration Skill (CLB). Creativity (CTV) shows an AVE value of 0.786, which is higher than its highest correlation of 0.554 to Critical Thinking Skill (CTC). Finally, Critical Thinking Skill (CTC) has an AVE value of 0.761, which is higher than its correlation with other variables, with the highest correlation of 0.583 with Communication Skill (CMC). The results of this analysis indicate that the Fornell-Larcker criterion in testing discriminant validity provides evidence that each construct is better at explaining variations in its own items than variations from other constructs, so the validity of the construct measurement model is acceptable.

The Heterotrait-Monotrait Ratio (HTMT) is a criterion analysis that must meet the requirement where the HTMT value should be greater than HTMT .85, i.e. 0.85 (Kline, 2011) or HTMT .90, i.e. 0.90 (Gold et al., 2001). As a statistical test, HTMT can test the null hypothesis (Ho: HTMT <1) vs (HA: HTMT ≥1) (Henseler et al., 2015; Sarstedt et al., 2020; Vinzi et al., 2010) with a 95% confidence interval HTMT that includes a value of 1 (i.e. HA), indicating the absence of discriminant validity. From Table 8, it can be seen that the HTMT values of the tested constructs meet the analysis criteria with values less than one. This indicates that the relationship between the constructs is very weak, which confirms the existence of discriminant validity for each construct tested. The diagonal represents the square root of the average variance extracted, while the other entries represent the correlation coefficient multiplied by two.

Table 8. HTMT Value

Variable	Collaboration Skill (CLB)	Communication Skill (CMC)	Coursework on Materials and changes (C)	Creativity (CTV)	Critical Thinking Skill (CTC)
Collaboration Skill (CLB)					
Communication Skill (CMC)	0.810				
Coursework on Materials and changes (C)	0.644	0.567			
Creativity (CTV)	0.631	0.756	0.766		
Critical Thinking Skill (CTC)	0.643	0.893	0.473	0.834	

Table 8 presents the results of the Heterotrait-Monotrait Ratio (HTMT) analysis used to test discriminant validity. This analysis requires that the HTMT value should be smaller than 0.85 or 0.90. As a statistical test, HTMT tests the null hypothesis (Ho: HTMT < 1) against the alternative hypothesis (HA: HTMT ≥ 1) with a 95% confidence interval. If the HTMT confidence interval includes a value of 1, then the alternative hypothesis is accepted, indicating the absence of discriminant validity.

In describing the model structure, we began by identifying the latent variables and their indicators. Latent variables are constructs that cannot be directly measured but are represented by multiple measurable indicators. We then established paths connecting the latent variables with their respective indicators, as well as paths linking one latent variable to another. Each path has a coefficient that indicates the strength and direction of the relationship. In the next step, we input the data into SmartPLS to estimate the model and examine the path coefficients, as well as the validity and reliability of the indicators. The results of this analysis were then visualized in a path diagram, which illustrates the relationships between variables and indicators, providing a clear representation of the model structure as show in Figure 1.

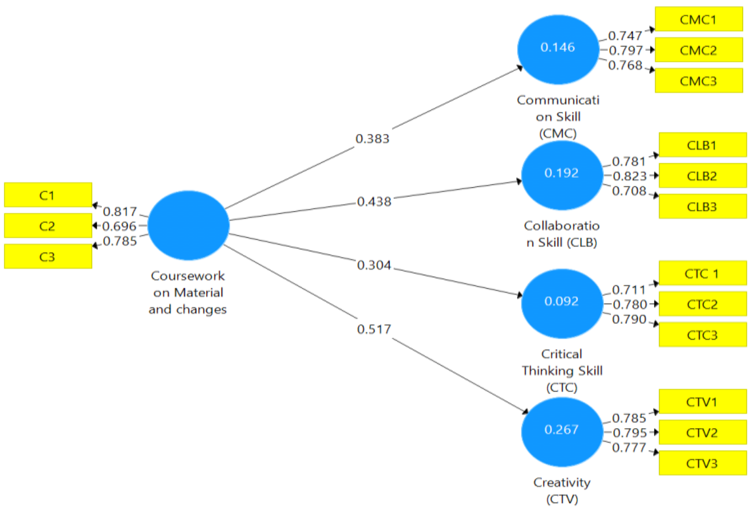

Figure 1. Structural Equation Model (SEM-PLS) Analysis of the Relationship Between Coursework on Material Changes and 4C Skills

The structural model depicted in Figure 1 illustrates the relationship between the independent variable "Coursework on Material and Changes" (C) and four key skills: Communication Skill (CMC), Collaboration Skill (CLB), Critical Thinking Skill (CTC), and Creativity (CTV). The independent variable is measured through three main indicators, namely C1, C2, and C3, with coefficients of 0.817, 0.826, and 0.785, respectively. "Coursework on Material and Changes" directly influences each of these skills, with the impact represented by the path coefficients linking the independent variable to each skill. For instance, the effect on Communication Skill (CMC) has a coefficient of 0.383, on Collaboration Skill (CLB) 0.438, on Critical Thinking Skill (CTC) 0.304, and on Creativity (CTV) 0.517.

Each skill is also measured by three indicators. Communication Skill (CMC) is measured by CMC1, CMC2, and CMC3, with coefficients of 0.747, 0.797, and 0.768, respectively. Collaboration Skill (CLB) is measured by CLB1, CLB2, and CLB3, with coefficients of 0.781, 0.823, and 0.708, respectively. Critical Thinking Skill (CTC) is measured by CTC1, CTC2, and CTC3, with coefficients of 0.714, 0.770, and 0.707, respectively. Creativity (CTV) is measured by CTV1, CTV2, and CTV3, with coefficients of 0.788, 0.755, and 0.777, respectively. This model demonstrates how "Coursework on Material and Changes" (C) influences students' skills in communication, collaboration, critical thinking, and creativity, as well as how each of these skills is measured through specific indicators.

Discussion

The choice of PLS-SEM analysis is also due to the existence of several other factors that can predict and explain the target construct, and can explore the relationship between constructs (Chin, 1998; Sarstedt et al., 2020). In addition, PLS-SEM is also able to analyze complex structural equation models consisting of many constructs and indicators (Sarstedt et al., 2020; Urbach et al., 2010). This study uses internal consistency reliability analysis to assess whether the items used can measure the same construct consistently in the score (Sarstedt et al., 2020; Vinzi et al., 2010). Internal consistency reliability is assessed using the composite reliability (CR) value. According to previous study CR values between 0.6 and 0.7 are acceptable for exploratory studies, while values between 0.7 and 0.9 are considered satisfactory for further studies. The CR value should be more than 0.7 to ensure adequate internal consistency (Gefen et al., 2000; Sarstedt et al., 2020). In addition, a Cronbach's Alpha value of more than 0.6 is also used as a measure of the reliability of items measuring a construct.

Convergent validity evaluation can be done through three tests, namely: 1) Outer Loading Assessment; 2) Composite Reliability (CR); and 3) Average Variance Extracted (AVE). First, the Outer Loading value must be more than 0.6 and the AVE value must be more than 0.5. Therefore, if in the initial analysis the Outer Loading value of one of the variables has a value between 0.4 to 0.7, it should be considered for removal if the removal of the item can increase the AVE or CR value. In addition, an Outer Loading value of more than 0.5 can also be considered because the item is considered to have good consistency (Chin, 1998; Hulland, 1999).

In this study, Partial Least Squares Structural Equation Modeling (PLS-SEM) was used to measure the Internal Consistency Reliability model, because it is able to analyze complex structural equation models consisting of many constructs and indicators (Sarstedt et al., 2020; Urbach et al., 2010). The choice of PLS-SEM was made because of its ability to predict and explain the targeted constructs. In addition, PLS-SEM also has advantages over other analytical methods. The findings of this study related to reliability are used to achieve overall consistency of results from items that measure the same construct, to assess whether the items used in the study can measure the same construct in the score value. Therefore, the results of this study consider the composite reliability (CR) value to assess the reliability of internal consistency with a CR value > 0.7 to ensure adequate internal consistency (Gefen et al., 2000; Sarstedt et al., 2020). In addition, a Cronbach's Alpha value of α > 0.6 was also considered to evaluate the reliability of the items measuring the constructs in this study. The findings showed that CR values for each construct ranged from 0.805 to 0.829, while Cronbach's Alpha values ranged from 0.636 to 0.690. This shows that the CR and Cronbach's Alpha values have met adequate standards, indicating that all 15 sub-constructs and formative constructs in this study have an appropriate level of internal consistency reliability (Gefen et al., 2000; Sarstedt et al., 2020).

This clearly shows that the 4C skills of prospective elementary school teachers can be explained by the constructs and sub-constructs in this study. To assess validity and reliability, CFA Convergent Validity analysis refers to how well an item can measure similar constructs in a study (Fitriati et al., 2023; Vinzi et al., 2010). Three tests are carried out in this convergent validity analysis, namely i) assessment of the output load > 0.5, ii) composite reliability (CR) value > 0.7, and iii) average extracted variance (AVE) value > 0.5. This research is based on the Outer Loading value > 0.5 used by Chin (1998) and Hulland (1999) which indicates that the items are considered as good estimates. Based on the analysis results, some items were eliminated based on the AVE value for each construct which must be greater than 0.5 (Bartlett et al., 2001) and the CR value which is greater than 0.7 (Sarstedt et al., 2020). The findings also show that all items meet the minimum requirement standard of Outer Loading. Thus, the results of this study clearly show that the items in this study have validity and reliability that meet the test requirements, which measure all elements, sub-constructs, and research agreements.

The instruments developed in this research can be used by lecturers to identify the extent to which prospective teachers' 4C (Critical thinking, Creativity, Communication and Collaboration) skills have developed during lectures. With this understanding, lecturers can design more effective learning activities, focusing on improving 4C skills according to the needs of each student. However, this research also has limitations, external factors such as personal motivation, educational background, and teaching experience of prospective teachers may influence the results of 4C skill development but were not measured in this research, so there is potential for bias in the research results.

4. CONCLUSION

The study demonstrates that the developed instrument successfully meets the validity and reliability criteria using the SEM-PLS analysis approach with SmartPLS. The core constructs—

communication skills, collaboration, critical thinking, creativity, and lecturer assignments—are conceptually sound, indicating strong construct validity and reliability. This instrument plays a crucial role in guiding the Indonesian Ministry of Education in curriculum design and policy alignment. By focusing on 4C skills, higher education curricula can be enhanced to better prepare graduates with the competencies demanded by the industry. The instrument also provides a valuable framework for future researchers in developing similar tools.

5. ACKNOWLEDGEMENT

We extend our deepest gratitude to the Ministry of Education, Culture, Research, and Technology for their support through the PMDSU research (grant number 051/E5/PG.02.00.PL.PMDSU/2024). This research would not have been possible without their generous funding and encouragement. Our heartfelt thanks also go to the Graduate School of Universitas Pendidikan Indonesia for their invaluable support and resources throughout this study. We are immensely grateful to all the respondents across Indonesia who participated in this research. Your contributions and insights have been indispensable to the success of this study. Thank you for your time and cooperation.

6. REFERENCES

Alismail, A. (2023). Teachers' perspectives of utilizing distance learning to support 21st century skill attainment for K-3 elementary students during the COVID-19 pandemic era. *Heliyon*, *9*(9), e19275. https://doi.org/10.1016/j.heliyon.2023.e19275.

Bartlett, J., Kotrlik, J. W., & Higgins, C. C. (2001). Determing appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, *19*(1), 43–50. https://banglajol.info/index.php/JBAU/article/view/74547.

Basaria, D., Zamralita, & Aryani, F. X. (2021). Peran Perfeksionisme Terhadap Prokrastinasi Akademik Pada Remaja di DKI Jakarta. *Jurnal Psibernetika*, *14*(1), 32–39. https://doi.org/10.30813/psibernetika.v.

Bialik, M., Martin, J., Mayo, M., & Trilling, B. (2016). Evolving Assessments for a 21st Century Education. In *Assessment Research Consortium*.

Bray, D. A., Girvan, D. C., & Chorcora, E. N. (2023). Students' perceptions of pedagogy for 21st century learning instrument (S-POP-21): Concept, validation, and initial results. *Thinking Skills and Creativity*, *49*(May 2022), 101319. https://doi.org/10.1016/j.tsc.2023.101319.

Chen, S., & Zhang, B. (2019). *Improving Prospective Teachers' Lesson Planning Knowledge and Skills through Lesson Study*. 549–575. https://doi.org/10.1007/978-3-030-04031-4_27.

Chin. (1998). Chin1998. In *MIS Quaterly* (Vol. 22, Issue 1, pp. vii–xvi).

Creswell, J. W. (2012). *Educational Research*. University of Nebraska.

Danks, N. P., Sharma, P. N., & Sarstedt, M. (2020). Model selection uncertainty and multimodel inference in partial least squares structural equation modeling (PLS-SEM. *Journal of Business Research*, *113*, 13–24. https://doi.org/10.1016/j.jbusres.2020.03.019.

Erdoğan, V. (2019). Integration of 4C competencies in online mathematics learning in junior high schools during the covid-19 pandemic. *International Journal of Education and Research*, *7*(11), 113–124. https://doi.org/10.1088/1742-6596/1918/4/042083.

Fitriati, F., Rosli, R., Iksan, Z., & Hidayat, A. (2023). Exploring challenges in preparing prospective teachers for teaching 4C skills in mathematics classroom: A school-university partnership perspectives. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2023.2286812.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312.

Gefen, D., Straub, D. W., Boudreau, M.-C., Gefen, D., Straub, D. W., & Boudreau, M. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. Communications of the association for information systems, 4(1), 7. *Communications of the Association for Information Systems*, *4*(1), 7. http://www.janrecker.com/quantitative-research-in-information-systems/.

Gold, A. H., Malhotra, A., & Segars, A. . (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems*, *34*(5), 688–698. https://doi.org/10.1002/ceat.201000522.

Griffin, P., & Care, E. (2015). Assessment and Teaching of 21st Century Skills. In *Assessment and Teaching of 21st Century Skills*. https://doi.org/10.1007/978-94-017-9395-7.

Hair, J. J. F., Hult, G. T. M., Ringle, M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM* (2nd ed.). Sage Publication.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8.

Hidayatullah, Z., Wilujeng, I., Nurhasanah, N., Gusemanto, T. G., & Makhrus, M. (2021). Synthesis of the 21st Century Skills (4C) Based Physics Education Research In Indonesia. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *6*(1), 88. https://doi.org/10.26737/jipf.v6i1.1889.

Hulland. (1999). Use Of Partial Least Squares (PLS) In Strategic Management Research: A Review Of Four Recent Studies. *Strategic Management Journal*, *20*, 195–204. https://doi.org/10.1038/aps.2012.31.

Khahro, S. H., & Javed, Y. (2022). Key Challenges in 21st Century Learning: A Way Forward towards Sustainable Higher Educational Institutions. *Sustainability (Switzerland)*, *14*(23). https://doi.org/10.3390/su142316080.

Koh, K., Delanoy, N., Bene, R., Thomas, C., Danysk, G., Hone, G., Turner, J., & Chapman, O. (2019). The Role of Authentic Assessment Tasks in Problem-Based Learning. *Postsecondary Learning and Teaching*, *3*, 17–24. https://doi.org/10.11575/pplt.v3i1.53144.

Laar, E. V, Deursen, A. J. A. M. V, Dijk, J. A. G. ., & Haan, J. D. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *SAGE Journal*, *10*(1), 1–14. https://doi.org/10.1177/2158244019900176.

Mardhotillah, N. I., & Yulhendri. (2022). Confirmatory factor analysis (cfa) interaksi dosen-mahasiswa dalam pembelajaran daring. *Jurnal Salingka Nagari*, *1*(1), 93–100. https://doi.org/10.24036/jsn.v1i1.13.

Maulana, Y., & Sopandi, W. (2022). Needs Analysis of Electronic Student Worksheets to Practice 4C Skills. *Jurnal Basicedu*, *6*(1), 602–611. https://www.academia.edu/download/83513644/pdf.pdf.

Maulana, Y., Wijayanti, I. E., & Solfarina, S. (2021). Development Electronic Student Worksheet based on Relating, Experiencing, Applying, Cooperating, and Transferring (REACT) in Introduction to Chemistry and Laboratory. *JTK (Jurnal Tadris Kimiya)*, *6*(2), 213–222. https://doi.org/10.15575/jtk.v6i2.14988.

Molla, A., Yayeh, M., & Bisaw, A. (2023). The current status of faculty members' pedagogical competence in developing 21st century skills at selected universities in Ethiopia. *Cogent Education*, *10*(2). https://doi.org/10.1080/2331186X.2023.2228995.

Nahar, S., Suhendri, Zailani, & Hardivizon. (2022). Improving Students' Collaboration Thinking Skill under the Implementation of the Quantum Teaching Model. *International Journal of Instruction*, *15*(3), 451–464. https://doi.org/10.29333/iji.2022.15325a.

Rodrigues, de S., Henrique, P., Reyes Junior, E., & del Corte Lora, V. (2024). Development, validation, and reliability of a measurement scale for investigating the principle of intercooperation. *Annals of Public and Cooperative Economics*, *95*(2), 553–572. https://doi.org/10.1111/apce.12426.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2020). Handbook of Market Research. In *Handbook of Market Research* (Issue July). https://doi.org/10.1007/978-3-319-05542-8.

Sunardi, S., & Doringin, F. (2020). The 4Cs Learning Model in Teacher Professional Development Program. *Humaniora*, *11*(2), 151–157. https://doi.org/10.21512/humaniora.v11i2.6508.

Trisnawati, W. W., & Sari, A. K. (2019). Integrasi Keterampilan Abad 21 Dalam Modul Sociolinguistics: Keterampilan 4C (Collaboration, Communication, Critical Thinking, Dan Creativity). *Jurnal Muara Pendidikan*, *4*(2), 455–466. https://doi.org/10.52060/mp.v4i2.179.

Urbach, F., Nils, & Ahlemann. (2010). Structural Equation Modeling in Information Systems Research Using Partial Least Squares. *Journal of Information Technology Theory and Application (JITTA)*, *11*(2), 5–40. https://repository.vlerick.com/handle/20.500.12127/6453.

Vinzi, V. E., Chin, W. W., Jörg, H., & Wang, H. (2010). PLS in Data Mining and Data Integration. In *Handbook of Partial Least Squares*. https://doi.org/10.1007/978-3-540-32827-8_16.

Widodo, S., & Wardani, R. K. (2020). Mengajarkan Keterampilan Abad 21 4C (Communication, Collaboration, Critical Thinking And Problem Solving, Creativity And Innovation) Di Sekolah Dasar. *MODELING: Jurnal Program Studi PGMI*, *7*(2), 185–197. https://doi.org/10.36835/modeling.v7i2.665.

Zhang, S., Du, J., Yue, H., Li, G., & Zhang, D. (2023). Study on the National Identity Education Intentions of Pre-Service Teachers: Evidence from PLS-SEM and fsQCA. *Sustainability (Switzerland)*, *15*(16). https://doi.org/10.3390/su151612633.

Zubaidah, S. (2018). Mengenal 4C: Learning and Innovation Skills untuk Menghadapi Era Revolusi Industri 4.0. *2nd Science Education National Conference*, *2*(2), 1–18. https://www.researchgate.net/profile/siti-zubaidah-7/publication/318013627.