Jurnal Penelitian dan Evaluasi Pendidikan
Volume 29, No. 1, June 2025 (58-70)
Online: http://journal.uny.ac.id/index.php/jpep

Confirmatory factor analysis of problem-based learning instrument for office management and business service program

Dika Putri Febriyanti[1]; Febrika Yogie Hermanto[1*]; Anthony Ogbolu Nnamdi[2]; Brillian Rosy[1]; Nova Septiana[1]
[1]Universitas Negeri Surabaya, Indonesia
[2]GCL Global Consulting, Nigeria
[*]Corresponding Author. E-mail: febrikahermanto@unesa.ac.id

ARTICLE INFO

Article History
Submitted:
08 April 2025
Revised:
23 April 2025
Accepted:
24 April 2025

Keywords
non-test instrument; confirmatory factor analysis; office management and business service program; problem-based learning

Scan Me:

ABSTRACT

This study examines the development and validation of a non-test instrument for assessing the Problem-Based Learning (PBL) model in vocational high schools, specifically within the Office Management and Business Services (OMBS) program. The instrument covers seven constructs: problem identification, fact identification, hypothesis formulation, self-directed learning, collaboration and cooperation, solution determination, and evaluation of the problem-solving process. This study uses a quantitative method with a Confirmatory Factor Analysis (CFA) approach to test the validity and reliability of the instrument. Using probability sampling with simple random sampling, 372 respondents (30 teachers and 342 students) from vocational schools in Surabaya participated. Data analysis was conducted using SPSS 16 for reliability testing and LISREL 8.8 for CFA. Results showed high internal consistency (Cronbach's Alpha = 0.949) and satisfactory anti-image coefficients (>0.5), indicating that the data met the requirements for factor analysis. The results of the model fit test also showed that the developed CFA with modification indices model met the criteria for a good fit with a Chi-square value of 170.51 with a degree of freedom (df) of 110, producing a p-value = 0.00019, an RMSEA value of 0.039 <0.08. The validity and reliability of the constructs in each instrument indicator also prove that this model meets a good fit with the resulting value above 0.5. Thus, this non-test instrument is worthy of being used as an assessment standard for evaluating problem-based learning in vocational high schools in the OMBS program to improve the quality of student learning outcomes.

To cite this article (in APA style):
Febriyanti, D. P., Hermanto, F. Y., Nnamdi, A. O., Rosy, B., Septiana, N. (2025). Confirmatory factor analysis of problem-based learning instrument for office management and business service program. *Jurnal Penelitian dan Evaluasi Pendidikan, 29(1)*, 58-70 doi: https://doi.org/10.21831/pep.v29i1.84368

INTRODUCTION

Education is the central pillar in developing quality human resources. 21st-century education demands developing critical thinking, problem-solving, and collaboration skills (Edwards, 2002). In this context, various innovative learning methods must be applied so students understand the theory and apply it in real situations following the Merdeka Curriculum. Problem-based Learning (PBL) is one of the learning methods that can be applied at various levels of education, including in Vocational High Schools. Problem-based Learning is designed to help students develop knowledge and problem-solving skills through social interaction (Salsabila & Muqowim, 2024). This concept is related to Lev Vygotsky's constructivism theory, which states that effective Learning utilizes social interaction to form new knowledge (Tohari & Rahman, 2024).

To effectively implement the Problem-Based Learning model, teachers require proper assessment that refers to a series of steps to determine specific aspects of an individual or group by collecting information on how students are well-achieving learning objectives (Aquino &

10.21831/pep.v29i1.84368

Yambi, 2020). In collecting assessment data, an assessment instrument is needed, where the assessment helps determine the quality of the learning process and measure student learning outcomes (Khasanah et al., 2020). The assessment instrument is used to collect information data from the assessment. One of the instruments in the assessment is a non-test instrument.

The non-test instrument is the collection of data or information on learning outcomes through cognitive (knowledge), affective (attitude) and psychomotor (soft skills) aspects (Dewi et al., 2024). The non-test instrument collects student assessment information, which can be a student performance assessment sheet. Student Performance Assessment is an approach to determine Learning progress by applying student knowledge and skills (Marlett, 2024). Performance assessment, according to (Pierce & O'Malley, 1992), is an assessment that can see the extent to which participants have mastered specific knowledge or skills related to the agreed continuum (rubric/criteria/standards).

Implementing the assessment sheet requires an assessment based on the Criterion-Reference Assessment (CRA) criteria. According to (Sternberg et al., 2022), CRA can be used to measure student performance against a series of predetermined standards (indicators) to improve teacher preparation in determining strategies and materials that can be used to improve student abilities. In the Problem-Based Learning model, the assessment sheet plays a role in assessing indicators (standards/criteria) that affect student learning outcomes, such as critical thinking skills, problem-solving, and student collaboration.

Based on observations in Vocational High School (VHS) in Surabaya, the data show that the Office Management and Business Services (OMBS) program has the potential to implement non-test instruments with the Problem-based Learning model. In OMBS, the program has administrative management skills, business communication, and customer service characteristics that require OMBS students to be involved in real-world problem-solving scenarios. So, in this case, the non-test assessment instrument in the Problem-Based Learning model is suitable to be implemented in the OMBS program. According to (Selviana & Puspasari, 2023), the OMBS program is appropriate and has the potential to involve students in critical thinking skills; collaborating with peers makes learning more student-centred. Then, some teachers of OMBS in Surabaya state that almost all teachers have implemented problem-based learning models. However, most teachers stated that they still have difficulty objectively determining the standard of non-test assessment instruments for the problem-based learning process. According to the Regulation of the Minister of Education and Culture No. 66 of 2013, the assessment standard in the learning process is essential and affects student learning outcomes. The Pranata et al. (2021) study revealed that assessments that dismiss to standard deviate from standards would give a biased capture of student knowledge or skill and selectivity in assessing student learning outcomes. Based on these observations, the gap between the needs in terms of assessment in the field and assessment instruments is still high, so it is necessary to develop non-test instruments for Problem-Based Learning. In order to meet the need for PBL non-test instruments, only the Septiana & Yogie (2024) study has developed standard non-test instrument instruments for OMBS teachers. They created seven indicators for non-test instruments in the Problem-based Learning model in the OMBS program. However, these indicators need to be tested more broadly for validity and reliability to evaluate learning outcomes in the OMBS program in Surabaya (Nurrahman et al., 2023).

Based on the situation, it is proven that the validity and reliability of non-test instruments directly affect the quality of education. Assessments that are not valid and reliable cannot be relied on to evaluate student abilities, hinder the achievement of learning objectives, and do not meet graduate competency standards (Kemendikbud, 2016). In implementing non-test instruments in the Problem-based Learning model that has been developed, an in-depth analysis is needed using Confirmatory Factor Analysis (CFA), which functions to test the relationship

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

between variables based on the hypothesized data model (Sarmento & Costa, 2019; Rusijono et al., 2020).

CFA also tests whether the indicators on the instrument meet good fit standards. If the analysis shows a valid and reliable relationship between variables, the non-test instrument developed can be relied on to reflect student learning outcome competencies. This study aims to ensure that the non-test instrument developed by Septiana & Yogie (2024) is suitable for implementation in the OMBS program. The results of the study are expected to improve the quality of valid and reliable non-test instruments, which can positively impact the evaluation of learning outcomes at OMBS Vocational High Schools in Surabaya.

RESEARCH METHOD

This study aims to validate that the non-test instrument that has been hypothesized by Septiana & Yogie (2024) has met the validity and reliability standards to be implied by teachers and students in implementing learning on the Problem-based Learning model at Office Management and Business Service (OMBS) program in Surabaya. This instrument consists of seven components, namely: (1) Problem Identification, (2) Fact Identification, (3) Formulating Hypotheses, (4) Self-Directed Learning, (5) Collaboration and Cooperation, (6) Determining Solutions to Problems, (7) Evaluation of the Problem Solving Process. This study used quantitative data with the Confirmatory Factor Analysis (CFA) approach. In line with Ghazali Nordin (2019) statement, the analysis of instruments and data through CFA has high reliability and validity (Putri & Febrilia., 2024). CFA is a statistical approach used to test the relationship structure between latent variables and their indicators; Timothy Brown's research in his book Confirmatory Factor Analysis for Applied Research (Brown, 2016). Confirmatory Factor Analysis (CFA) is a statistical model of factor analysis used to test the extent to which data is by the hypothesized factor model (Sureshchandar, 2023).

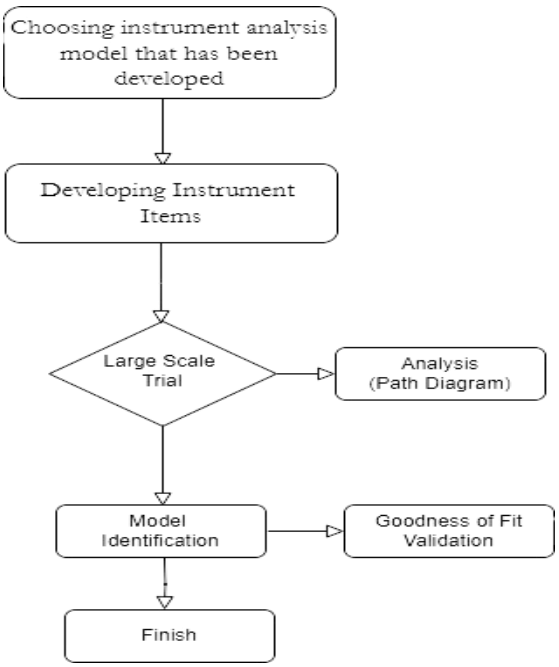

Figure 1. The procedure of validating non-test instruments

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

In this study, the instrument that has been developed is a non-test instrument with a Likert scale of 1-5, so the questionnaire developed by the researcher uses a Likert scale of 1-5 according to the hypothesis. The indicators in each of these statements have a Likert scale model with 1 (Strongly Disagree), 2 (Disagree), 3 (Neutral), 4 (Agree), and 5 (Strongly Agree) based on (Croasmun, 2011; Nurrahman et al., 2022). The population and sample of the study are students and teachers of Vocational High Schools in the OMBS program in Surabaya who have taken learning through the Problem-based Learning model. Figure 1 shows a flow diagram of the validation procedure for the assessment instrument in this study.

The first step is to select a non-test instrument previously developed in the study. Using a non-test instrument that has been developed is more effective because it has undergone a validation stage. The selected non-test instrument was developed by Septiana and Yogie (2024) to measure the effectiveness and learning outcomes using the Problem-based Learning (PBL) model OMBS program. In this study, the instrument is an indicator that can evaluate aspects of the Problem-based Learning model.

The second step is to develop question items based on the indicators set according to the Problem-based Learning model. In this stage, the assessment (score) of the items on each indicator developed uses a Likert scale with five response categories: Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree. The development of these instrument items aims to ensure that each question item can effectively measure the aspects needed in the Problem-based Learning model based on the perspectives of Teachers and Students of the OMBS program in Surabaya.

After the instrument items were developed, the third step was to conduct a small-scale trial. In this study, a small-scale trial involved 30 respondents, 10 teachers, and 20 students of Surabaya's Office Management and Business Service (OMBS) program. This small-scale trial aims to evaluate the instrument items' clarity, readability, and relevance before they are widely used. The selection criteria for educational units used as subjects in this study were Teachers and Students of State and Private Vocational High Schools that have the OMBS program in Surabaya and have undergone PBL learning. In a small-scale trial using SPSS and EFA analysis, it was proven that the items developed were reliable, with a score of 0.878. This is intended so that the research results reflect the diversity of conditions and characteristics of educational units in the OMBS environment and can be continued to the next stage, namely, large-scale distribution.

The fourth step is to conduct a large-scale trial. This trial aims to determine the extent to which the instruments that have been developed can be used effectively in measuring aspects of the Problem-based Learning model. This study uses Probability Sampling with the Simple Random Sampling method to ensure a good population representation (Sugiyono, 2019). This study uses an unlimited population, so researchers cannot determine the total population (Asrulla et al., 2023). So, a total of 372 respondents were involved in the large-scale trial. The large-scale trial uses Google Forms for data collection, consisting of 30 teachers and 342 Surabaya Office Management and Business Service (OMBS) program students.

The data obtained from the large-scale trial were last analysed using a path diagram with the Confirmatory Factor Analysis (CFA) technique. At this stage, researchers can see the relationship between the variables measured in the assessment instrument through a multidimensional design. The software used in this multidimensional CFA analysis is LISREL 8.8, which makes it easier for researchers to see how the established indicators correlate.

Following the generation of the path diagram, the next step involved testing the construct validity and reliability. This process is crucial for confirming whether the resulting measurement model meets the minimum required values. Construct validity was assessed by examining loading factor scores > 0.5, while construct reliability was determined by calculating CR and AVE scores. Subsequently, the model's goodness-of-fit was evaluated using LISREL 8.8, which

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

indicates the degree to which the model can be considered a good fit. Suppose the results of the CFA construct validity and reliability tests and the model's goodness-of-fit test meet the established good-fit standards. In that case, the developed instrument can be deemed valid and suitable for implementation. Conversely, revisions to the instrument or model are necessary if the model fails to meet the fitness criteria.

FINDINGS AND DISCUSSION

The Confirmatory Factor Analysis on non-test instruments in the problem-based learning model began with the statement of the Criterion Reference Assessment Standard in the Regulation of the Minister of Education and Culture No. 66 of 2013, which states that assessment standards are used as a process of collecting information to determine student learning outcomes (Kemendikbud, 2016; Sugiyono, 2019). Based on the statement of the assessment standard criteria, researchers utilized the non-test instrument developed by Septiana & Yogie (2024) to analyze more deeply the validity, reliability, and relationship between variables. Testing the assessment instrument is one way to show that the instrument that has been developed is of quality by proving its validity (Pada et al., 2018). The Confirmatory Factor Analysis on non-test instruments in the problem-based learning model began with the statement of the Criterion Reference Assessment Standard in the Regulation of the Minister of Education and Culture No. 66 of 2013, which states that assessment standards are used as a process of collecting information to determine student learning outcomes. Based on these assessment standards, researchers utilized previously developed non-test instruments to conduct an in-depth analysis regarding the suitability between theoretical models and empirical data and to identify whether the available indicators were relevant to conditions in the field through Confirmatory Factor Analysis (Bariyyah et al., 2024).

There are seven constructs developed for non-test instruments in the Problem-based Learning model, including (1) Problem Identification, (2) Fact Identification, (3) Formulating Hypotheses, (4) Self-Directed Learning, (5) Collaboration and Cooperation, (6) Determining Solutions to Problems, (7) Evaluation of the Problem-Solving Process. Furthermore, a reference is formulated to determine the construction of the developed instrument to produce a total of 18 indicators. Then, these indicators are developed into 18 statements that are considered capable of measuring each construct of the non-test instrument in the Problem-based Learning model for teachers and students of Vocational High School OMBS in Surabaya. In writing the instrument items, the researcher focuses on what needs (aspects) are needed in evaluating student learning outcomes through the Problem-based Learning model. Problem-based learning is a learning approach that emphasizes the active involvement of students in integrating theory and practice to solve real problems collaboratively (Savery, 2006).

In the field of research, particularly in the validation of indicators in assessment instruments, the concept of reliability is crucial. It ensures that these instruments, which have been hypothesized, meet established reliability standards and can be applied effectively. Reliability refers to the degree to which an instrument yields consistent results when used repeatedly. Reliability in this study was proven using the Alpha-Cronbach reliability coefficient. Reliability testing in this study also used SPSS 16 software. Based on the reference (Imaduddin et al., 2022), it is known that a construct or variable is reliable if Cronbach's alpha value is > 0.60. The reliability test results of this study using SPSS 16 software with Cronbach's alpha can be seen in Table 1.

According to Table 1, the reliability score of Cronbach's alpha for the non-test instrument is 0.949, where Imaduddin et al. (2022) states that the score category of the non-test instrument in this study possesses a high level of reliability. Then, Sainuddin et al. (2022) state that the high-level reliability test can be reflected in consistent measures of the indicators. Therefore, these

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

results demonstrate that the non-test instrument for the problem-based learning model is indeed reliable (Ghazali & Nordin, 2019).

Table 1. Blueprint of the Creative Thinking Skills test

Reliability Statiscs	
Cronbach's Alpha	N of Items
0.949	18

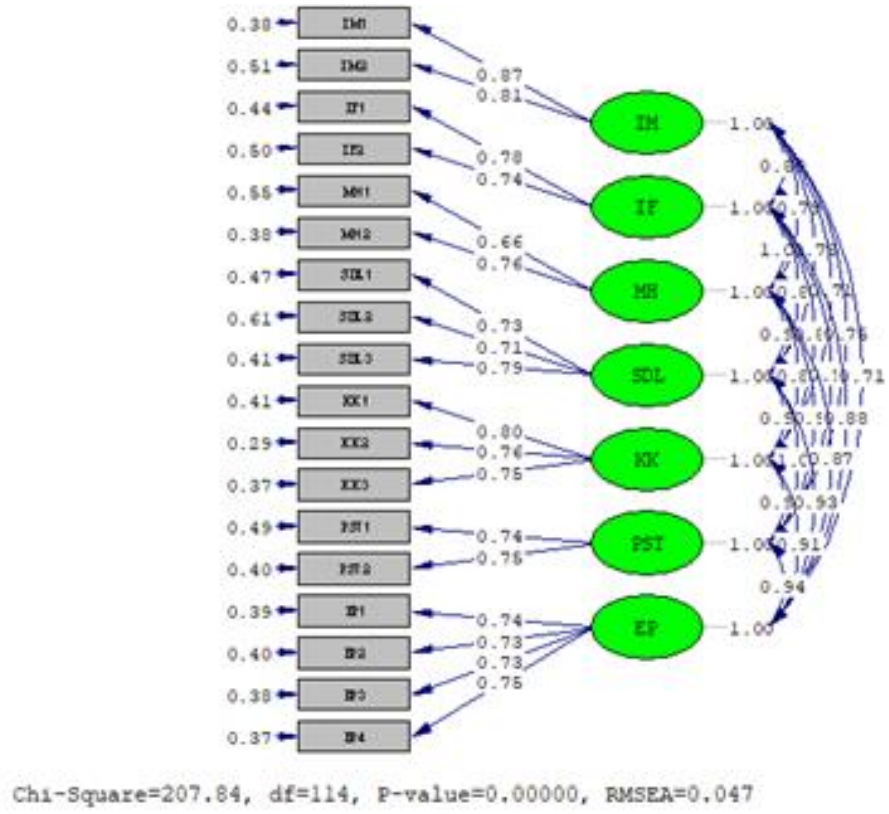

Figure 2. CFA analysis result multidimensional

Figure 2 shows the results of the multidimensional CFA analysis with a Chi-square value of 207.84, a degree of freedom (df) of 114, a p-value of 0.0000, and an RMSEA of 0,047. In a CFA study, researchers might consider modifications when initial results indicate a suboptimal model fit. CFA can help detect incorrect model specifications, thus allowing for statistical and theoretical model improvements that can improve the quality of the instrument through Modification Indices (Byrne, 2016).

Based on the results, the p-value was detected as 0.0000, so modification indices needed to be carried out through the indicators MH2 - KK2, SDL1 - KK1, SDL2 - KK2, and KK3 - PST2, which had a reduction value on Chi-Square (Decrease in Chi-Square) and an addition to the p-value (see Figure 3). Although the p-value was still not met after the modification indices, there was still an increase in value; this often happens if Chi-square (X^2) has a large sample size. Which states that the Chi-square value is less than 2(pdf) based on (Jöreskog & Sörbom, 1993; Nurrahman et al., 2023), resulting in a p-value = 0.00019, which does not meet the criteria> 0.05. The RMSEA value of 0.039 <0.08, according to de Jonge (2006), indicates a good model

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

fit. In addition, the loading factor value of each item shows a value> 0.4, which means that all items are accepted (Herwin & Nurhayati, 2021). Therefore, based on the RMSEA and loading factor values, it can be concluded that the developed model is fit.

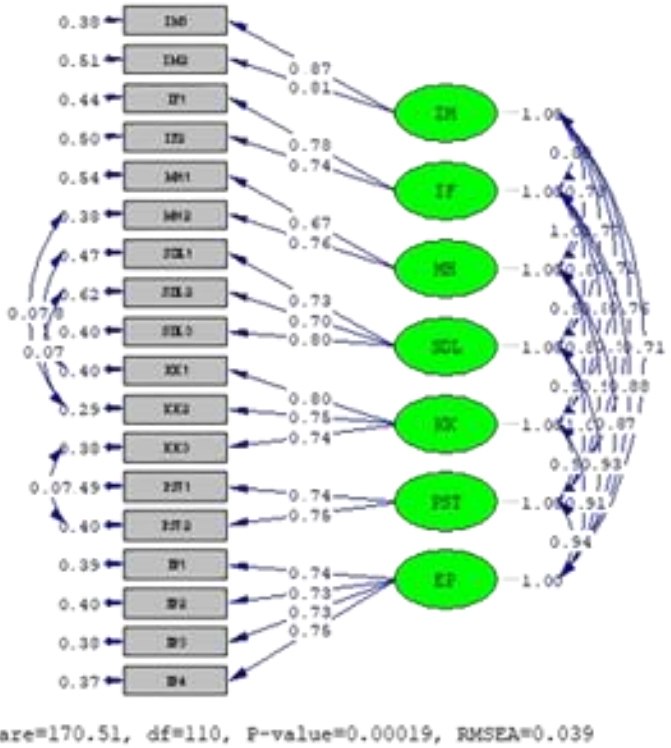

Figure 3. CFA analysis result multidimensional after modification

Table 2 presents details of the results of the non-test instruments that have been developed to measure learning outcomes through the Problem-based Learning model, which are distributed into seven constructs and consist of 18 items. The development of these items is based on previous research and is aligned with the operational definition of each construct. The assessment instrument developed (Septiana & Yogie, 2024) shows strong construct validity with each item having a loading factor that meets the minimum standard of > 0.5 (Nurbaiti, 2021), namely with a value range between 0.66 to 0.87, which proves that there is an indication of a strong relationship between the item (indicator) and the construct being measured.

Then, the analysis shows that the non-test instrument that has been developed has met the criteria for Construct Reliability and Average Variance Extracted (AVE) based on Nurbaiti (2021). The construct reliability score is reliable, with the range of 0.6 ≤ CR ≥ 0.7, while the Average Variance Extracted (AVE) score is more than 0.5; the data can be seen in Table 3. The analysis shows that all constructs of the non-test instrument have met the assessment standards, confirming the consistency of the constructs and items being measured. Specifically, the construct that measures the Problem-solving Process Evaluation (EP) variable shows the highest construct reliability with a score of 0.8496, indicating an excellent level of internal consistency among the items that form the construct. Conversely, the construct that measures the Hypothesis Formulation (MH) variable has the lowest construct reliability, which is 0.6897, which, although it meets the minimum requirements, indicates that some items may need to be reviewed to improve consistency.

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

Table 2. Loading factor problem based learning model instruments

No.	Aspects	Items	Factor Loading
1	Problem Identification (IM)	Accuracy of Problem Identification	0.87
		Accuracy of Problem Formulation	0.81
2	Identification of Facts (IF)	Accuracy of Fact Identification	0.78
		Accuracy in using information/Reference Source	0.74
3	Constructing a Hypothesis (MH)	Accuracy of Hypothesis Formulation from Problem Formulation	0.67
		Accuracy of Problem-solving Strategies	0.76
4	Self Directed Learning (SDL)	Initiative in Providing Solutions to Problems	0.73
		Thinking Independently to Determine Problem Solving Strategies	0.70
		Responsibility for Resolving Problems Raised	0.80
5	Collaboration and cooperation (KK)	Active Interaction with Group Members in Solving Problems	0.80
		Collaboration with Group Members in Solving Problems	0.75
		Collaboration with Group Members in Solving Problems	0.74
6	Determining Solutions to Problems (PST)	Structured Problem Solving	0.74
		Accuracy of Delivery of Ideas or Problem Solving Solutions	0.75
7	Problem Solving Process Evaluation (EP)	Assessment/Evaluation of Problem Solving Solutions	0.74
		Assessment of the Responses to Each Problem Solving Alternative	0.73
		Reflection of Problem Solving Solutions	0.73
		Accuracy of Conclusion Drawing from Problem Solving Discussion Results	0.75

Table 3. Construct reliability of non-test instruments

Variable	CR	AVE	Desc.
IM	0.7602	0.6135	Reliable
IF	0.7108	0.5515	Reliable
MH	0.6897	0.5273	Reliable
SDL	0.7666	0.5233	Reliable
KK	0.8293	0.6185	Reliable
PST	0.7138	0.5550	Reliable
EP	0.8496	0.5855	Reliable

In terms of Average Variance Extracted (AVE), which measures the variance of indicators explained by the latent construct, the construct measuring the Collaboration and Cooperation

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

(KK) variable showed the highest value of 0.6185, indicating that this construct explains most of the variance of its indicators. Meanwhile, the construct measuring the Self-Directed Learning (SDL) variable also has the lowest AVE value, which is 0.5233, which, although still above the threshold of 0.5, this construct may require further improvement to increase convergent validity. Overall, the results of this analysis confirm that the assessment instrument that has been developed has met the requirements for construct reliability, allowing researchers to use this instrument to be implied in evaluating learning outcomes in the Problem-based Learning model.

In the model fit test, some measurements can be used as benchmarks for a model to be declared 'fit', namely the Absolute Fit Measure and the Incremental Fit Measure. In this study, the value that will measure the Absolute Fit Measure is proven by Chi-square ($\mathcal{X}^2$) p-value, Goodness of Fit Index (GFI), and Root Mean Square Error of Approximation (RMSEA). In contrast, the Incremental Fit Measure will be proven by the Non-Normed Fit Index (NNFI), Adjusted Goodness Fit Index (AGFI), and Comparative Fit Index (CFI) values. It can be seen that 6 of the seven fit index criteria have been met. Namely, the Goodness of Fit Index (GFI), Adjusted Goodness Fit Index (AGFI), Comparative Fit Index (CFI), Non-Normed Fit Index (NNFI) and Root Mean Square Error of Approximation (RMSEA). The data can be seen in Table 4.

Although the Chi-square ($\mathcal{X}^2$)/DF value is ≤ 2, the p-value does not meet the recommended criteria; this often occurs when Chi-square ($\mathcal{X}^2$) has a large sample size (Khairi et al., 2021). Therefore, the Root Mean Square Error of Approximation (RMSEA) value of less than 0.08, accompanied by a Goodness of Fit Index (GFI) value of more than 0.9, can be used to state that the model fits (de Jonge, 2006; Umar & Nisa, 2020). Based on this, the model in this study is considered fit because the RMSEA value is less than 0.08, which is 0.047, with a GFI value of more than 0.9, which is 0.94. So, it means that the construct of the non-test instrument developed in the Problem-based Learning model consisting of 7 constructs has been proven valid and significant for measuring learning outcome assessment because it meets at least one fit criterion in the Absolute Fit Measure and Incremental Fit Measure. In addition, the 18 items contained in the instrument have also been proven valid and significant for each aspect, as evidenced by the Standardized Loading Factor score of more than 0.5.

Table 4. Suitability of the first-order confirmatory factor analysis model

No	Criteria	Cut off Value	Results	Desc
1	CMIN(Chi square)/DF	≤ 2	170.51 / 114 = 1.49	Fit
	p-value	>0.05	0.00019	Not fit
2	NNFI	≥ 0.9	0,99	Fit
3	GFI	≥ 0.90	0.95	Fit
4	RMSEA	≤ 0.08	0.039	Fit
5	AGFI	≥ 0.90	0.92	Fit
6	CFI	≥ 0.9	0.99	Fit

Learning in the Merdeka curriculum is based on student-centred learning, which must be implemented using the PBL model. This learning requires a non-test instrument that can evaluate students' skills in critical thinking, problem-solving, collaboration, and social skills comprehensively (Arta, 2024). This shows the necessity of measurable validity and reliability for the instruments employed, ensuring their capacity to capture learning outcomes and bolster the effectiveness of problem-based learning within the OMBS program. Integrating non-test assessment instruments in a learning model with assessment aspects on critical thinking, communication, social, and collaboration skills in PBL can improve student learning outcomes in vocational high schools (Fatihatussa'adah et al., 2024).

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

This study provides practical benefits for teachers, especially in implementing Problem-Based Learning in the OMBS program. Validated non-test assessment instruments can be standardized tools for teachers to evaluate student learning outcomes using critical thinking, collaboration, communication, and problem-solving skills (Savery, 2006). This instrument enables teachers to conduct more objective and quantifiable assessments. Implementing non-test assessment instruments within a problem-based learning model aligns with the principles of the Merdeka curriculum. This allows for a more accurate recording of OMBS students' scholastic abilities, enabling the resulting scores to reflect their actual competencies more precisely.

Furthermore, this study's results are based on respondents from teachers and students in the OMBS program in Surabaya. Hence, the interpretation and application of the findings are still contextual and limited to the needs of the Problem-Based Learning assessment model of the OMBS program in Surabaya. Thus, the generalizability of the developed instrument needs to be considered carefully because the needs of problem-based learning (PBL) assessment in Surabaya may differ from those in other areas with diverse student characteristics, resources, and curriculum.

CONCLUSION

This study produces a Non-Test Instrument in the problem-based learning model that has been proven valid and reliable to be applied by teachers and students of the OMBS Program in Surabaya. First, constructing the non-test instrument in the Problem-based Learning model that has been developed involves seven constructs: problem identification, fact identification, formulating hypotheses, self-directed learning, collaboration and cooperation, determining solutions to problems, and evaluating the problem-solving process. Then, these aspects are developed into indicators that are adjusted to the context of the needs of teachers and students when evaluating the learning process using the problem-based learning model. The final product produced is 18 statement items with a frequency scale with details of 2 items for the first construct, two items for the second construct, two items for the third construct, three items for the fourth construct, three items for the fifth construct, two items for the sixth aspect, and four items for the seventh construct. Second, the quality of the developed instrument indicators has been proven by Cronbach's alpha reliability with SPSS 16, construct validity with CFA, and goodness of fit model suitability analysis through Lisrel 8.8. This demonstrates that the instrument indicators prepared can be used to measure student learning outcomes through the problem-based learning model of the Vocational High School of the OMBS Program in Surabaya. Thus, according to the indicators developed, non-test instruments accurately measure critical thinking, problem-solving, collaboration, and social skills.

Furthermore, the provision of measurable non-test instruments needs to be prepared by teachers by providing appropriate training in creating or using instruments for learning in schools. This training is important so teachers can create or adjust the context of learning implemented in the OMBS program. In addition, further research is also carried out by developing non-test instruments on Project-Based Learning or other learning that is often implemented in schools so the reflections of student abilities can be read according to the scores obtained in subjects at school, especially on instruments that have been validated and reliable.

ACKNOWLEDGMENT

We would like to thank Universitas Negeri Surabaya and all participants (students and teachers of the OMBS Program in Surabaya) in this study. Hopefully, this study can contribute to increasing the quality of education, particularly in non-test instrument development.

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

Conflict of interests

There are no known conflicts of interest associated with this publication.

REFERENCES

Aquino, T. De, & Yambi, C. (2020). *Assessment and evaluation in education*. July.

Arta, G. Y. (2024). Asesmen dalam pendidikan: Konsep, pendekatan, prinsip, jenis, dan fungsi. *Jurnal Pendidikan, Bahasa dan Budaya, 3*(3), 170–190. https://doi.org/10.55606/jpbb.v3i3.3925

Asrulla, A., Risnita, R., Jailani, M. S., & Jeka, F. (2023). Populasi dan sampling (kuantitatif), serta pemilihan informan kunci (kualitatif) dalam pendekatan praktis. *Jurnal Pendidikan Tambusai, 7*(3), 26320–26332. https://doi.org/10.31004/jptam.v7i3.10836

Bariyyah, K., Pratama, A. M., Tsaqofah, A., Pratiwi, N. P., Siwi, R. A. C., Ahfadzi, Y. F., Budiyono, A. L., & Maizura, N. (2024). Confirmatory factor analysis (CFA) untuk validasi instrumen sexual bullying questionnaire (SEBUQU). *Jurnal EDUCATIO: Jurnal Pendidikan Indonesia*, *10*(2), 258–265. https://doi.org/10.29210/1202424887

Brown, T. A. (2016). *Confirmatory Factor Analysis for Applied Research* (second, Vol. 4, Issue 1).

Byrne, B. M. (2016). *Structural Equation Modeling with Amos*.

Croasmun, J. T., & Ostrom, L. (2011). Using Likert-type scales in the social sciences. *Journal of Adult Education, 40*(1), 19–22. https://files.eric.ed.gov/fulltext/EJ961998.pdf

de Jonge, J. (2006). Multivariate data analysis (Sixth Edition). *Gedrag & Organisatie, 19*(3), 123–145. https://doi.org/10.5117/2006.019.003.007

Dewi, I. K., Arsurahmah, A., Syaharani, D., Putri, S., & Marhadi, H. (2024). Tinjauan terhadap efektivitas instrumen penilaian non tes di sekolah dasar: Studi kasus Kelas VI SDN 130 Pekanbaru. *Jurnal Ilmiah Kajian Multidisipliner*, *8*(5), 538–545.

Edwards, R. (2002). Lifelong learning, lifelong learning, lifelong learning: A recurrent education? In *Lifelong learning: Education across the lifespan* (pp. 1–13). Routledge. https://doi.org/10.4324/9780203183328-1

Fatihatussa'adah, I., Yamtinah, S., Ariani, S. R. D., Wiyarsi, A., Widarti, H. R., Shidiq, A. S., & Abrori, F. M. (2024). Fostering collaboration and enhancing student learning achievement through the integration of ethnoscience in the Common Knowledge Construction Model with podcast media. *Indonesian Journal on Learning and Advanced Education, 6*(3), 295–314. https://doi.org/10.23917/ijolae.v6i3.23222

Ghazali, N., & Nordin, M. S. (2019). Measuring meaningful learning experience: Confirmatory factor analysis. *International Journal of Innovation, Creativity and Change*, *9*(12), 283–296.

Herwin, & Nurhayati, R. (2021). Measuring students' curiosity character using confirmatory factor analysis. *European Journal of Educational Research*, *10*(2), 773–783. https://doi.org/10.12973/eu-jer.10.2.773

Imaduddin, M. F., Maulani, H., & Taufik, I. H. (2022). Test the validity and reliability of Arabic learning questions. *Arabi : Journal of Arabic Studies*, *7*(2), 198–207. https://doi.org/10.24865/ajas.v7i2.523

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Scientific Software International.

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

Kemendikbud. (2016). *Permendikbud_Nomor 23_Tahun 2016*. BPK RI. https://peraturan.bpk.go.id/Details/224434/permendikbud-no-23-tahun-2016

Khairi, M. I., Susanti, D., & Sukono, S. (2021). Study on structural equation modeling for analyzing data. *International Journal of Ethno-Sciences and Education Research*, *1*(3), 52–60. https://doi.org/10.46336/ijeer.v1i3.295

Khasanah, M. M., Ramli, M., & Dwiastuti, S. (2020). Developing a dynamic assessment instrument to assess reasoning skills about bacteria. *Jurnal Penelitian dan Evaluasi Pendidikan*, *24*(1). 1-12. https://doi.org/10.21831/pep.v24i1.28815

Marlett, D. (2024). *Effective Performance-Based Assessments Empower Students to Demonstrate Knowledge and Skills*. Leraning-Focused. https://learningfocused.com/effective-performance-based-assessments-empower-students-to-demonstrate-knowledge-and-skills/

Nurbaiti, B. (2021). *Tutorial Pengolahan Data dengan Lisrel* (Issue April).

Nurrahman, A., Cahyani, M. D., Nurfatmawati, L., Wibowo, H., & Analysis, C. F. (2023). *Developing E-Learning Evaluation Instruments : A Study in Vocational*. *3*(3), 163–174.

Nurrahman, A., Sukirno, S., Pratiwi, D. S., Iskandar, J., Rahim, A., & Rahmaini, I. S. (2022). Developing student social attitude self-assessment instruments: A study in vocational high school. *REID (Research and Evaluation in Education)*, *8*(1), 1–12. https://doi.org/10.21831/reid.v8i1.45100

Pada, A. U. T., Mustakim, S. S., & Subali, B. (2018). Construct validity of creative thinking skills instrument for biology student teachers in the subject of human physiology. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(2), 119–129. https://doi.org/10.21831/pep.v22i2.22369

Pierce, L. V., & O'Malley, J. M. (1992). *Performance and portfolio assessment for language minority students*. Program Information Guide Series, 9. National Clearinghouse for Bilingual Education. https://eric.ed.gov/?id=ED346747

Pranata, A. H., Sumarno, & Dwijayanti, I. (2021). Analisis kesulitan penyusunan instrumen penilaian bagi guru sekolah dasar di Gugus Dewi Sartika Kecamatan Reban. *Widya Sari*, *23*(5), 105–110.

Putri, R. D. P., & Febrilia, Y. (2024). Confirmatory factor analysis (CFA) on the quality of critical thinking instruments, motivation, and learning achievement of learners. *Pendagogia: Jurnal Pendidikan Dasar,* *4*(2), 73–84. https://jurnal.educ3.org/index.php/pendagogia/article/view/183.

Rusijono, R., Jaedun, A., Kartowagiran, B., Ahman, A., Laliyo, L. A. R., & Mam, S. (2020). Developing the teacher's social competency assessment instrument in the fourth industrial revolution era. *Jurnal Penelitian dan Evaluasi Pendidikan*, *24*(2), 125–135. https://doi.org/10.21831/pep.v24i2.29482.

Sainuddin, S., Subali, B., Jailani, & Elvira, M. (2022). The development and validation of prospective mathematics teachers' holistic assessment tools. *Ingénierie des Systèmes d'Information*, *27*(2), 171–184. https://doi.org/10.18280/isi.270201

Salsabila, Y. R., & Muqowim, M. (2024). Korelasi antara teori belajar konstruktivisme Lev Vygotsky dengan model pembelajaran problem based learning (PBL). *LEARNING: Jurnal Inovasi Penelitian Pendidikan dan Pembelajaran*, *4*(3), 813–827. https://doi.org/10.51878/learning.v4i3.3185

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v29i1.84368

Sarmento, R. P., & Costa, V. (2019). *Confirmatory factor analysis – A case study*. arXiv. https://doi.org/10.48550/arXiv.1905.05598

Savery, J. R. (2006). Overview of problem-based learning: Definitions and distinctions. *Interdisciplinary Journal of Problem-Based Learning*, *1*(1), 9–20. https://doi.org/10.7771/1541-5015.1002

Selviana, M., & Puspasari, D. (2023). Analyzing the application of the problem-based learning model in office management and business services subjects. *Journal of Office Administration: Education and Practice*, *3*(2), 84–94. https://doi.org/10.26740/joaep.v3n2.p84-94

Septiana, N., & Hermanto, F. Y. (2024). *Construct validity of non-test instrument for problem-based learning method. Journal of Office Administration: Education and Practice*, *4*(1), 1–9. https://doi.org/10.26740/joaep.v4n1.p1-9

Sternberg, R. J., Chowkase, A., Parra-Martinez, F. A., & Landy, J. (2022). Criterion-referenced assessment of intelligence as adaptation to the environment: Is it possible, plausible, or practical? *Journal of Intelligence*, *10*(3), 57. https://doi.org/10.3390/jintelligence10030057

Sugiyono. (2019). *Metode Penelitian Kuantitatif Kualitatif dan R&D* (Sutopo (ed.)). Alfabeta.

Sureshchandar, G. S. (2023). Quality 4.0 – a measurement model using the confirmatory factor analysis (CFA) approach. *International Journal of Quality and Reliability Management*, *40*(1), 280–303. https://doi.org/10.1108/IJQRM-06-2021-0172

Tohari, B., & Rahman, A. (2024). Konstruktivisme Lev Semonovich Vygotsky dan Jerome Bruner: Model pembelajaran aktif dalam pengembangan kemampuan kognitif anak. *Nusantara: Jurnal Pendidikan Indonesia*, *4*(1), 209–228. https://doi.org/10.14421/njpi.2024.v4i1-13.

Umar, J., & Nisa, Y. F. (2020). Uji validitas konstruk dengan CFA dan pelaporannya. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia (JP3I)*, *9*(2), 1–11. https://doi.org/10.15408/jp3i.v9i2.16964

ISSN (print) 2685-7111 | ISSN (online) 2338-6061