Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

A PRELIMINARY CONFIRMATORY FACTOR ANALYSIS OF A MATHEMATICS LEARNING INDEPENDENCE INSTRUMENT FOR ELEMENTARY SCHOOL STUDENTS

Ahmad[1)]
Mariano Dos Santos[2)]
Habib Ratu Perwira Negara[3)]
Johri Sabaryati[4)]
[1,3)]Computer Science, Bumigora University, Mataram, Indonesia
[2)] Istituto Católico para a Formação de Professores (ICFP) Baucau, Timor-Leste
[3)]Study Program, Faculty, University, City, State
[4)]Study Program in Physics Education, Muhammadiyah University, Mataram, Indonesia
e-mail: ahmad@universitasbumigora.ac.id

ABSTRACT

This study aims to assess the validity of the construct of the mathematics learning independence instrument for elementary school students using Confirmatory Factor Analysis (CFA). The development of the instrument began with the preparation of an initial draft of 20 items representing five theoretical constructs: learning initiative, self-regulation, confidence, perseverance, and learning responsibility. After going through the stages of content validation by experts and readability screening, a total of 12 items were retained for empirical testing. Data were collected through a quantitative survey design of 69 elementary school students and analyzed using a lavaan package in the R software with a Maximum Likelihood (ML) estimate as the baseline analysis. The results of the model fit evaluation showed that the 5-factor 12-item model had a less than ideal match rate (χ^2 /df = 4.19; CFI = 0.624; TLI = 0.436; RMSEA = 0.216; SRMR = 0.142). In contrast, Factors 3, 4, and 5 resulted in inaccurate parameter estimates (Heywood's case). Overall, this instrument is not psychometrically acceptable and requires structural revision, addition of indicators, and retesting using ordinal estimators (Weighted Least Squares Mean and Variances Adjusted / WLSMV) in larger samples.

Keywords: Validity of construction; Analysis of confirmatory factors; Independence from mathematics learning.

INTRODUCTION

Basic education has a fundamental role in forming the foundation of thinking skills, attitudes, and values that will determine students' success at the next level of education (German & Randel, 1998; Inusah et al., 2023; Sousa-Ferreira et al., 2021). In the context of mathematics learning, logical, systematic, and independent thinking skills are competencies that need to be developed from an early age (Ahmad et al., 2021; Rawani et al., 2024). However, the reality on the ground shows that most elementary school students are still heavily dependent on teachers in the learning process (Jansen in de Wal et al., 2020; Rusydiyah et al., 2020). This can be seen from the lack of student initiative in completing math tasks, unpreparedness to face learning difficulties, and lack of confidence in designing problem-solving strategies. This condition shows that the independence of mathematics learning for elementary school students still needs to be improved through learning oriented towards the development of independent learning skills.

Independent learning is one of the skills of the 21st century that is the focus of educational transformation in Indonesia (Dalas et al., 2020; Maimun & Bahtiar, 2023; Mulyono, 2017). In the Independent Curriculum, learning independence is included as part of the Pancasila Student Profile which emphasizes students' ability to take responsibility for their own learning processes and outcomes (Nurhayati et al., 2022). In the context of mathematics learning, learning independence not only means students' freedom in determining how they learn, but also includes the ability to strategize, monitor progress, and evaluate learning outcomes. Thus, students who have high learning independence will be able to understand mathematical concepts in depth, solve problems with confidence, and adapt to a variety of challenging learning situations.

Various previous studies have shown that learning independence has a strong relationship with math learning achievement. Zimmerman (2000) Calling learning independence a part of independent learning, which is an active process in which students consciously set goals, strategies, and evaluations in learning. Zimmerman (2000) and Pintrich (2004) emphasizing that students who are able to control behavior and motivation in learning will show better academic achievement. However, in the context of basic education, especially in fields such as Central Lombok, these abilities are often not developed optimally because learning is still teacher-centered. Teachers tend to be the main source of knowledge, while students act as passive recipients. As a result, the aspect of learning independence is not the focus of competency development in mathematics learning (Bellibaş & Gümüş, 2023; Matsuyama et al., 2019).

Efforts to encourage independence in mathematics learning need to be supported by instruments that are able to measure the level of student independence validly and reliably (Ahmad et al., 2025). This measurement is important so that educators can get an objective picture of the extent to which students have independent learning abilities and what factors still need to be developed (Andriatna et al., 2024; Barros et al., 2022; Duell et al., 2015). However, the instruments available so far have generally been developed in the context of secondary or tertiary education, so they do not fully correspond to the cognitive and affective characteristics of elementary school students. On the other hand, the socio-cultural context in an area such as Central Lombok also influences the way students perceive learning and their responsibility to the learning process. This requires instruments that are contextually developed and tested for validity in the primary education environment in the area.

In this study, a mathematics learning independence instrument was developed to measure five main dimensions that represent important aspects of student learning independence, namely learning initiative, self-regulation, self-confidence, perseverance, and learning responsibility. These five dimensions are governed by the self-regulated learning theory proposed by Zimmerman (2000) and contextualized with the characteristics of elementary school students in Central Lombok. This instrument consists of 20 statements compiled using a four-point Likert scale, ranging from "strongly disagree" to "strongly agree". The development of this instrument not only aims to measure the level of students' mathematics learning independence, but also provides an evaluation tool that can be used by teachers and researchers in identifying learning needs that support the formation of learning independence.

In the testing of psychological or educational instruments, the validity of the construct is a very important aspect because it shows the extent to which the item in the instrument actually measures the theoretical construct in question. According to Hair et al. (2020), the validity of the construct encompasses two main components, namely convergent validity and discriminant validity. Convergent validity indicates the extent to which indicators in a single construct are positively correlated with each

other, while discriminant validity ensures that each construct is empirically different from each other. One widely used approach to assessing the validity of a construct is Confirmatory Factor Analysis (CFA) (Nu'man et al., 2021; Thompson, 2004). The CFA allows researchers to test the compatibility between theoretical models and empirical data, so that it can be used to assess the structural strength of factors in an instrument.

Confirmatory factor analysis is not only useful for verifying the structure of theoretically established factors, but also helps to determine which indicators are most representative for each dimension of the construction (Herwin & Nurhayati, 2021; Kolenikov, 2009; Schreiber et al., 2006). In the context of this study, the CFA was used to assess the extent to which the five-factor model of mathematics learning independence was developed based on data obtained from elementary school students. This approach also provides important information about the validity and reliability of the instrument, including the mean of extracted variance (AVE) as a measure of convergent validity, and composite reliability (ρc) as a measure of construction reliability (Dos Santos & Cirillo, 2023; Valentini & Damásio, 2016). Thus, the results of the CFA analysis are expected to provide a strong empirical basis for the development of mathematics learning independence instruments at the elementary school level.

Research on the validity of mathematics learning independence instruments at the elementary school level is still relatively limited, especially in Indonesia and more specifically in the Central Lombok region. Most existing studies still focus on secondary school or college students, while research on elementary school students often stops at the descriptive stage. In addition, previous studies generally used exploratory factor analysis, which was only able to identify patterns of relationships between indicators without confirming their theoretical structure. Therefore, this study is important because it uses confirmatory factor analysis to test the validity of instrument constructions developed based on theory and local context.

The urgency of this research also lies in its contribution to the development of educational measurement tools that are in accordance with the characteristics of elementary school students in the region. By using a rigorous psychometric approach, the results of this study are expected to provide a valid and reliable instrument to measure students' mathematics learning independence. This empirically validated instrument can be used by teachers, researchers, and education policymakers to conduct more targeted evaluation and learning interventions, especially to encourage independent learning attitudes in elementary school students in Central Lombok.

This study aims to test the validity of the construction of mathematics learning independence instruments in elementary school students using Confirmatory Factor Analysis (CFA). This research is expected to make a theoretical contribution to the development of the concept of learning independence in elementary education, as well as make a practical contribution in the form of standard instruments that can be used in the context of mathematics learning in elementary schools. Thus, the results of this study not only strengthen the theoretical foundation regarding the importance of independent learning in mathematics learning, but also become the first step in efforts to improve the quality of character-based and competency-based learning in the 21st century.

METHOD

This study is a preliminary validation study using a quantitative approach of Confirmatory Factor Analysis (CFA) to test the validity of the construct of the mathematics learning independence instrument of elementary school students in Central Lombok Regency. The draft instrument tested consisted of

12 statement items distributed on five latent factors: learning initiative (2 items), self-regulation (4 items), confidence (2 items), perseverance (2 items), and learning responsibility (2 items). Each item was measured using a four-point Likert scale (1 = "Strongly Disagree" to 4 = "Strongly Agree"). Although the structure of two indicators per factor meets the minimum mathematical identification limit in a multivariate (just-identified) model, methodologically the two-factor factor is statistically fragile, especially when tested on a limited pilot sample size (N = 69).

The research data was collected from a limited sample of 69 elementary school students selected through targeted sampling. The use of this limited sample size was intended as an initial exploratory step to refine the item and test the feasibility of the instrument before conducting large-scale testing in the future. The data collection process was carried out by filling out a questionnaire directly in the classroom with the help of researchers and subject teachers. Before the CFA analysis was conducted, the data were first examined to ensure their statistical feasibility, including testing for normality, data integrity, and consistency of responses.

Data analysis was carried out using lavaan packets in the R software. Model fit evaluation was carried out through several goodness-of-fit indices, including Chi-Square (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR)(Bentler, 1990; Schubert et al., 2017; Shi et al., 2018). Model estimation at this initial exploration stage was performed using the Maximum Likelihood (ML) method as a baseline analysis. The selection of ML in this pilot trial was based on normative considerations to detect parameter instability sensitivity and variance residues in a limited sample (N = 69). Although some literature notes that 4-category ordinal data can be tested with ML as an initial continuous approach, methodologically the use of ML at the ordinal scale risks violating the assumption of multivariate normality, underestimating standard errors, and triggering the Heywood case. Therefore, the use of robust estimators specifically for categorical ordinal data, namely Weighted Least Squares Mean and Variances Adjusted (WLSMV), is recognized as theoretically more appropriate and is absolutely recommended for reestimating the model in subsequent analyses.

This approach is used to evaluate the contribution of each indicator to the measured construct, as well as to obtain preliminary empirical evidence regarding the latent structure of the instrument. The results of this pilot CFA analysis are expected to be the basis for evaluation to revise the instrument so that it is better ready to be tested on a larger sample and representative of a wider population.

RESULTS AND DISCUSSION

Before evaluating the individual parameters of each item, the overall fit of the model was evaluated to determine the extent to which the 12-item 5-factor structure corresponded to the empirical data. The results of the evaluation showed that the model had a poor fit: $\chi^2 = 184.52$ (df = 44, $p < 0.001$), $\chi^2/df = 4.19$ (≤ threshold 3.0), CFI = 0.624 (≥ threshold 0.90), TLI = 0.436 (≥ threshold 0.90), RMSEA = 0.216 (≤ threshold 0.08), and SRMR = 0.142 (≤ threshold 0.08). These indices, which far exceed these standard criteria, confirm that the hypothesized five-factor model as a whole is not yet qualified and requires structural revision.

Model and Specification Factor

The results of the following Confirmatory Factor Analysis (CFA) analysis present the structure of the measurement model consisting of five latent factors with a number of observed indicators. Each factor is formed based on a theoretical construction that has been developed in the previous stage.

This model is estimated using the Maximum Likelihood (ML) method with the help of *lavaan* packets on the R software.

Table 1. Specification of Conconfirmatory Factor Analysis Model

Factors	Forming Items	Number of Items
Factor 1	Item_3, Item_10	2
Factor 2	Item_9, Item_13, Item_7, Item_4	4
Factor 3	Item_19, Item_11	2
Factor 4	Item_12, Item_17	2
Factor 5	Item_14, Item_20	2
Estimation Method	Maximum Likelihood (ML)	
Number of Respondents	69	

Based on the results of the model specification in Table 1, this instrument consists of five latent factors that represent the main dimensions of learning independence, namely learning initiative, self-regulation, confidence, perseverance, and learning responsibility. Each factor was measured with two to four indicators adapted from the theory of self-regulated learning and has gone through a content validation process by experts. The structure of these five factors was estimated using the Maximum Likelihood (ML) method in 69 elementary school student respondents selected through targeted sampling. The use of ML estimators in this initial exploratory analysis served to identify the baseline behavior of the measurement model. However, it is important to affirm as a major methodological limitation that the use of ML on categorically limited ordinal data (4-point Likert scale) with a pilot sample size (N = 69) is highly susceptible to violating multivariate assumptions about normality. This violation of categorical assumptions is the main cause of parameter estimation instability and the emergence of estimation errors (Heywood cases), which reinforces the urgency of using ordinal estimators such as WLSMV in subsequent reanalysis and validation.

The CFA model used describes the relationship between latent constructs and observed indicators to ascertain whether each statement item in the instrument truly represents the aspect in question. The first factor, consisting of Item_3 and Item_10, is assumed to represent aspects of learning initiatives that reflect the student's ability to initiate and direct his or her learning activities independently. The second factor formed by Item_9, Item_13, Item_7, and Item_4 describes the self-regulation dimensions associated with students' ability to manage time, strategies, and learning resources. The third factor, consisting of Item_19 and Item_11, indicates confidence in learning math, while the fourth factor with Item_12 and Item_17 indicates the student's perseverance in the face of learning difficulties. Meanwhile, the fifth factor formed by Item_14 and Item_20 relates to students' responsibility for their learning process and outcomes.

Judging from the model specifications, four of the five latent factors in this instrument were measured with only two indicators. Statistically, although two indicators are the minimum limit requirements for a latent factor to be mathematically identified in a multivariate CFA model, the two indicator factor is statistically fragile and has great limitations. The use of two indicators per factor is susceptible to instability of parameter estimation, inability to detect specific measurement errors, and is a major trigger for the emergence of improper solutions (Heywood cases), especially when combined with the limited sample size of the pilot (N = 69). Therefore, the current model structure is not recommended to be maintained. For later stage validation, it is highly recommended to increase the

number of conceptually relevant indicators (at least 3 to 4 items per factor) or to combine overlapping latent dimensions to improve model stability and identification. This model represents a framework of thought that the independence of mathematics learning is not a single construct, but rather consists of several psychological dimensions that interact with each other (Hajjina & Retnawati, 2022).

Model Estimation Results

Once the measurement model is established, the next stage is to estimate the parameters to determine the contribution of each indicator to the latent construct being measured. This analysis results in the values of the load factor, standard load, and determination coefficient (R-Square) for each item. These values show the extent to which the indicator can represent the factors formed. The results of the CFA model estimation based on empirical data can be seen in Table 2 below.

Table 2. Model Estimation Results (Load Factor and Variance)

Factors	Item	Forecast	Std.lv	Std.all	R-Square
Factor 1	Item_3	1.000	0.953	0.998	0.996
	Item_10	0.297	0.284	0.292	0.085
Factor 2	Item_9	1.000	0.307	0.356	0.127
	Item_13	1.162	0.357	0.386	0.149
	Item_7	1.755	0.539	0.564	0.319
	Item_4	0.650	0.199	0.199	0.040
Factor 3	Item_19	1.000	20.766	25.067	N/A
	Item_11	0.000	0.006	0.007	0.000
Factor 4	Item_12	1.000	21.046	20.698	N/A
	Item_17	-0.000	-0.004	-0.004	0.000
Factor 5	Item_14	1.000	0.014	0.015	0.000
	Item_20	828.536	11.863	12.818	N/A

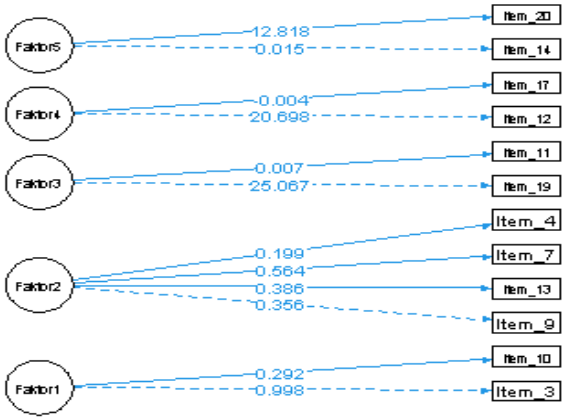

Figure 1. **Path Chart of CFA Analysis Results**

The estimated results of the Confirmatory Factor Analysis (CFA) model as shown in Table 2 and Figure 1 show the strength of the relationship between the observed indicators and the hypothetical latent constructs in the mathematics learning independence instruments of elementary school students. In general, the model consists of five latent factors, each measured by two to four indicators. The

estimation values, standard load (Std.all), and R-Square are used to assess the contribution of each indicator to the construct it represents. The high value of the load factor indicates that the indicator has a good ability to explain the variance of latent constructions.

In Factor 1, Item_3 showed a very high load value (Std.all = 0.998) with an R-Square of 0.996. This indicates that the indicator is able to explain more than 99% of the construction variance, thus being a strong representation of the dimensions of the learning initiative. In contrast, Item_10 has a low load value (Std.all = 0.292) with an R-Square of 0.085, which indicates a relatively small contribution to the construction. Although still at the threshold of minimal acceptance, these results indicate the need to re-evaluate the item, both in terms of editorial and its relevance to the theoretical dimension.

In Factor 2, which represents self-regulation, the load value of the indicator ranges from 0.199 to 0.564. Item_7 have the highest contribution (Std.all = 0.564, R-Square = 0.319), indicating that this indicator is the most capable of explaining the construct compared to the other three items. While Item_4 has the lowest contribution (Std.all = 0.199, R-Square = 0.040), which indicates that this item is less representative of the self-regulation dimension. Overall, although Factor 2 shows a positive charge pattern, the magnitude of the variation between the indicators suggests that some items need to be refined in order for the construct to be more consistent.

In contrast to the previous two factors, Factor 3, Factor 4, and Factor 5 exhibit an unstable estimation pattern that results in even extreme values. For example, in Factor 3, Item_19 has a standard load value of 25,067, while Item_11 is only 0.007. Similarly, in Factor 4, Item_12 has a load value of 20.698, while Item_17 is close to zero (-0.004). In contrast to the first two factors, Factor 3, Factor 4, and Factor 5 show unreasonable estimation results (improper solutions). The standard factor load (λ) at Item_19 (λ = 25.067), Item_12 (λ = 20.698), and Item_20 (λ = 12.818) is statistically invalid because the standard load value should ideally not exceed 1.0. This phenomenon indicates the emergence of the Heywood Case, which is triggered by negative error variance due to the limitation of the pilot sample size (N = 69) and the statistically very fragile (fragile) two-indicator factor structure. In a psychometric context, this can occur due to limited sample sizes, very high indicator correlation, or model specification errors.

The visualization in Figure 1 shows the relationship between the factors and the indicators with the direction of the arrows pointing to each measurement item. The different lengths and thicknesses of the lines illustrate the variation in the strength of the relationship between the variables. It can be seen that Factor 1 and Factor 2 have a more proportional relationship direction than the other three factors, suggesting that the two factors are relatively more empirically stable. Meanwhile, the long lines with extreme load values in Factor 3, Factor 4, and Factor 5 show an imbalance in the relationship, confirming that the model structure has not yet achieved optimal conformity. Overall, the results of the CFA model estimation show that some indicators, particularly in Factor 1 and Factor 2, show a consistent and significant relationship with their construction, while other factors still need to be corrected by the model. These findings suggest that although some of the dimensions of the instrument are quite valid, the overall structure of the model cannot be considered stable (Polit & Beck, 2006; Wahyuni, 2023).

Construction Validity (AVE)

Once the *factor load value* is obtained, the next step is to test the validity of the construct through the calculation of the Extracted Mean Variance (AVE). The AVE value indicates the proportion of

indicator variance that the latent construct successfully explained. The greater the AVE value, the better the convergent validity of the construct. The results of the AVE calculation for each factor in the CFA model are presented in Table 3 below.

Table 3. Construct Validity (Extracted Mean Variance – AVE)

Factors	AVE Value	Information
Factor 1	0.532	Meet minimum criteria (≥ 0.5)
Factor 2	0.157	Invalid
Factor 3	Uninterpretable	Improper solution / The Heywood case
Factor 4	Uninterpretable	Improper solution / The Heywood case
Factor 5	Uninterpretable	Improper solution / The Heywood case

The construct validity test is performed to assess the extent to which the indicators in each factor are able to represent the latent construct being measured. This test is carried out through the calculation of the Average Extracted Variance (AVE) which shows the proportion of the variance of the indicator that can be explained by the latent construct compared to the variance error. In general, an AVE value greater than or equal to 0.50 indicates that the construct has good convergent validity, while a value below 0.50 indicates that the indicator in the construct is not strong enough to explain the latent variable (Valentini & Damásio, 2016).

Based on the results of the analysis in Table 3, Factor 1 numerically meets the minimum threshold of convergent validity with an AVE value of 0.532. This value shows that approximately 53.2% of the variance of the forming indicator can be explained by the latent variable. This result is in line with the findings in Table 2, where Item_3 has a dominant load factor (0.998). However, given the overall fit of the CFA model (overall model fit) shows a poor level of fit (CFI = 0.624; TLI = 0.436; RMSEA = 0.216; SRMR = 0.142), the AVE value in Factor 1 is not interpreted as the validity of an established or fully tested construct. Factor 1 (learning initiative) is only an initial indicator that shows relatively better performance than the other dimensions, but still requires structural respecification and retesting.

In contrast, Factor 2 shows a very low AVE value, i.e. 0.157, so it does not meet the criteria for convergent validity. This value indicates that only about 15.7% of the indicator's variance can be explained by latent constructions, while most of the variance comes from measurement errors. This condition is in line with the load factor results in Table 2, where some items in Factor 2 (such as Item_4 and Item_9) have relatively low load values. This suggests that the indicators on this factor are inconsistent in measuring self-regulating dimensions and need to be corrected, either by revising the redaction of the item or adding new indicators that are more conceptually relevant.

Meanwhile, the AVE calculation of Factor 3, Factor 4, and Factor 5 did not produce a valid psychometric estimate due to the emergence of a negative error variance (Heywood cases). Therefore, the AVE values for these three factors could not be interpreted statistically. This phenomenon of parameter convergence failure confirms that the indicators in Factors 3, 4, and 5 fail to measure the hypothetical latent constructs of these pilot samples, thus requiring a respecification of the model structure. These findings confirm that all three factors fail to estimate the pilot sample, so the indicators cannot be claimed to have convergent validity. These values far exceed the limits of statistical fairness of the CFA, indicating model inconsistencies or misestimates (known as the Heywood case). This phenomenon usually arises due to correlations between indicators that are too high, small sample sizes, or models that do not match the data. In the context of this study, the results confirm that the

three factors do not yet have sufficient empirical stability to represent the expected latent construction. Thus, the interpretation of extreme AVE values needs to be done with caution and cannot be used as a basis for assessing the validity of the construct directly (Sping, 2009).

Composite Reliability (ρC)

In addition to the validity of the construction, the reliability of each factor also needs to be tested to ensure internal consistency between the indicators in a single construction. Construction reliability is calculated using Composite Reliability (ρC) which reflects the extent to which these indicators are able to provide stable measurement results. A high reliability value indicates that the construction has good internal consistency. The results of the construction reliability calculation are presented in Table 4 below.

Table 4. Composite Reliability (ρC)

Factors	Value ρC	Information
Factor 1	0.638	Meet minimum criteria (≥ 0.6)
Factor 2	0.393	Unreliable
Factor 3	Uninterpretable	Improper solution / The Heywood case
Factor 4	Uninterpretable	Improper solution / The Heywood case
Factor 5	Uninterpretable	Improper solution / The Heywood case

A construction reliability test is performed to assess the internal consistency between indicators in a single latent construct. High reliability indicates that all indicators in a factor provide stable and consistent measurement results for the constructed being measured. In this study, construction reliability was calculated using Composite Reliability (ρC), which is considered more accurate than Alpha Cronbach because it takes into account the weight factor of each indicator. A ρC value of ≥ 0.60 is generally used as the minimum limit to declare that a construction is reliable (Hair et al., 2019).

Based on the results of the analysis in Table 4, Factor 1 produced a ρC value of 0.638, which is technically above the minimum reliability limit (0.60). This value indicates a relatively adequate initial internal consistency between Item_3 and Item_10 in measuring the dimensions of the learning initiative. This finding is in line with the AVE results (0.532). However, because the overall model fit is very poor (CFI = 0.624, RMSEA = 0.216), this level of reliability is still tentative and cannot be used as a reference that the construction fit has been psychometrically stable.≥

On the other hand, Factor 2 indicates a low reliability value, which is 0.393, so it does not meet the minimum limit of reliability. This value indicates that the relationship between indicators in self-regulation construction is still weak and inconsistent. This condition can be caused by differences in students' understanding of statement items or incompatibility of indicators with the learning context at the elementary school level. Therefore, it is necessary to re-examine items that have low load factors, such as Item_4 and Item_9, to improve the internal strength of the construction.

In line with the AVE results, the Composite Reliability values of ρC) for Factor 3, Factor 4, and Factor 5 also cannot be interpreted due to statistical estimation failures (Heywood cases). The absence of valid reliability values in these three factors confirms that the dimensions of confidence, perseverance, and learning responsibility cannot be measured consistently by the indicators currently used. This failure confirms that the dimensions of confidence, perseverance, and learning responsibility cannot be measured consistently using the current indicators. Cases like this usually occur when the

model experiences non-convergence errors or factor variance to negative (Heywood cases). This condition is reinforced by previous findings in the construction validity test (AVE) which also resulted in extreme values on all three factors. Thus, very high ρC values on these factors cannot be interpreted as high reliability, but rather as an indication that the model requires structural modification.

Overall, the results of the construction reliability test showed that only Factor 1 showed a relatively maintainable initial reliability indicator, while the other factors experienced estimation failures. However, due to the poor overall model fit index, the positive tendency in Factor 1 was not enough to conclude that the dimension was absolutely valid and reliable. The implication of these findings is that the entire measurement structure of the mathematics learning independence instrument including Factor 1 still requires editorial revision, structural modification, and further validation in a larger sample. This improvement process is important so that this instrument is not only theoretically valid, but also empirically reliable in measuring the learning independence of elementary school students consistently in various learning contexts (Surucu & Maslakci, 2020).

Discriminatory validity

The final step in the CFA analysis is to test the validity of the discriminators to ensure that each construct has a clear difference from each other. This test is performed using the Heterotrait-Monotrait Ratio (HTMT), which measures the degree of correlation between latent constructs. HTMT values smaller than 0.85 indicate that these constructs are empirically different from each other. The results of the inter-factor discrimination validity test are presented in Table 5 below.

Table 5. Discriminant Validity (HTMT Ratio)

	Factor1	Factor2	Factor3	Factor4	Factor5
Factor1	1.000				
Factor2	0.270	1.000			
Factor3	0.283	0.870	1.000		
Factor4	0.587	0.727	0.632	1.000	
Factor5	0.356	0.415	0.140	**1.585**	1.000

Based on the results in Table 5, discriminant validity testing using the Heterotrait-Monotrait ratio (HTMT) showed discrimination problems in several pairs of factors. First, the relationship between Factor 2 (self-regulation) and Factor 3 (self-confidence) resulted in an HTMT value = 0.870. Because this value exceeded the conservative threshold of 0.85, there are early indications of conceptual overlap between the two dimensions. This indicates that elementary school students in this sample tend not to be able to clearly distinguish between behavior of managing learning strategies and confidence in learning mathematics.

Second, very serious violations of discriminatory validity occurred in the pairs of Factor 4 (diligence) and Factor 5 (learning responsibility) with an HTMT value = 1.585, which far exceeded the maximum limit of 0.85. HTMT values above 1.0 confirmed the occurrence of extreme multicollinearity and empirical discrimination failure. The indicators in Factor 4 and Factor 5 were not conceptually differentiated by respondents and measured essentially the same construct.

Overall, the results of the discriminant validity test showed that the instrument experienced significant construct differentiation constraints, both indications of slight overlap in the Factor 2 and Factor 3 pairs (HTMT = 0.870) and severe discrimination failures in the Factor 4 and Factor 5 pairs

(HTMT = 1.585). These findings reinforce the evidence that the current five-factor structure suffers from misspecification errors (model misspecification) and requires the addition or reduction of dimensions in future revisions of the instrument. In addition, a follow-up validation process can be carried out using a respecification model or confirmatory composite analysis (CCA) to ensure that the remaining constructs are truly independent (Hair et al., 2020; Hubona et al., 2021; Schamberger et al., 2023).

CONCLUSION

Based on the results of the Confirmatory Factor Analysis (CFA), the mathematics learning independence instrument for elementary school students as a whole has not been psychometrically acceptable because it produces a poor model fit index (CFI = 0.624; TLI = 0.436; RMSEA = 0.216; SRMR = 0.142). Of the five dimensions tested, Factor 1 (learning initiative) showed only a relatively better initial indicator than the other factors in terms of convergent validity and composite reliability. However, given the overall poor fit of the model, the validity and reliability of Factor 1 cannot be interpreted as an established construct. The other dimensions experienced model instability, discriminant validity problems, and the emergence of the Heywood case. Therefore, further research is recommended to reestimate the model using appropriate estimators for ordinal categorical data such as WLSMV, revise or remove weak indicators, increase the sample size (N > 69), and test the structure of the respecification factor before the instrument can be widely applied in elementary schools.

REFERENCE

Ahmad, A., Dos Santos, M., Helmiyatun, H. (2025). Analysis of exploratory factors for the validity of mathematics learning independence instruments. *El Midad: Journal of PGMI Department 17*(1), 34–52.

Ahmad, A., Parihin, P., Halimatuzzahrah, H., Jannah, M., Fadli, H., Rismayati, R., & Suriyati, S. (2021). Meningkatkan motivasi belajar siswa melalui pemberian reward dan punisment bagi siswa Madrasah Tsanawiyah. Jurnal Riset Teknologi dan Inovasi Pendidikan, 4(2), 267–278.

Andriatna, R., Sujadi, I., Budiyono, B., Kurniawati, I., Wulandari, A. N., & Puteri, H. A. (2024). Junior high school students' numeracy in geometry and measurement content: Evidence from the minimum competency assessment result. *AIP Conference Proceedings*, 3046(1), 020036. https://doi.org/10.1063/5.0194570

Barros, A., Santos, H., & Santos-Silva, F. (2022). A systematic review of health literacy measurement instruments in Portugal. *Portuguese Journal of Public Health*, *40*(3), 172–187. https://doi.org/10.1159/000525890

Bellibaş, M. Ş., & Gümüş, S. (2023). The effect of learning-centred leadership and teacher trust on teacher professional learning: Evidence from a centralised education system. *Professional Development in Education*, *49*(5). https://doi.org/10.1080/19415257.2021.1879234

Bentler, P. M. (1990). Comparative fit index in structural models. *Bulletin of Psychology, 107*(2). https://doi.org/10.1037/0033-2909.107.2.238

Dalas, N., Hadiyanto, & Muhaimin. (2020). The relationship between learning independence, family support, facilities, and institutional services through student learning quality. *Journal of Critical Reviews* 7(6), 815-821. https://doi.org/10.31838/jcr.07.06.140

Dos Santos, P. M., & Cirillo, M. Â. (2023). Construction of an extracted average variance index for construction validation in a structural equation model with adaptive regression. *Communication in Statistics: Simulation and Computing*, *52*(4). https://doi.org/10.1080/03610918.2021.1888122

Duell, P., Wright, D., Renzaho, A. M. N., & Bhattacharya, D. (2015). Optimal health literacy measurement for the clinical setting: A systematic review. *Patient Education and Counseling*, *98*(11), 1295–1307. https://doi.org/10.1016/j.pec.2015.04.003

German, T., & Randel, J. (1998). Basic education. In The Reality of Aid 1998–1999: An independent review of poverty reduction and development assistance, 11, 27–34. https://doi.org/10.4324/9781315067070-6

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis (7th ed.).* Pearson Prentice Hall.

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*, 101-110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report PLS-SEM results. *European Business Review* 31(1), 2-24. https://doi.org/10.1108/EBR-11-2018-0203

Hajjina, C. Y. N., & Retnawati, H. (2022). Analysis of confirmatory factors to measure self-efficacy and outcome expectations in mathematics learning. *AlphaMath : Journal of Mathematics Education*, *8*(2), 111-120. https://doi.org/10.30595/alphamath.v8i2.12604

Herwin, H., & Nurhayati, R. (2021). Measuring students' curiosity character using confirmation factor analysis. *European Journal of Education Research*, *10*(2) 773-783. https://doi.org/10.12973/EU-JER.10.2.773

Hubona, G. S., Schuberth, F., & Henseler, J. (2021). A clarification of confirmatory composite (CCA) analysis. *International Journal of Information Management*, *61*. https://doi.org/10.1016/j.ijinfomgt.2021.102399

Inusah, F., Missah, Y. M., Najim, U., & Twum, F. (2023). Agile neurological systems for managing elementary education. *Intelligent Systems with Applications*, *17*. https://doi.org/10.1016/j.iswa.2023.200178

Jansen in de Wal, J., van den Beemt, A., Martens, R. L., & den Brok, P. J. (2020). The relationship between job demands, work resources, and teacher professional learning: is it explained by self-determination theory? *Studies in Continuing Education*, *42*(1) 17-39. https://doi.org/10.1080/0158037X.2018.1520697

Kolenikov, S. (2009). Analysis of confirmatory factors using konfa. *Stata Journal*, *9*(3) 329-373. https://doi.org/10.1177/1536867x0900900301

Maimun, M., & Bahtiar, B. (2023). Students' learning independence and critical thinking skills using mobile learning technology. *European Journal of Education Research*, *12*(4) 1815-1830. https://doi.org/10.12973/eu-jer.12.4.1815

Matsuyama, Y., Nakaya, M., Okazaki, H., Lebowitz, A. J., Leppink, J., & Van der Vleuten, C. (2019). Does changing from a teacher-centered to a learner-centered context promote self-regulated learning: A qualitative study in a Japanese undergraduate setting. *BMC Medical Education*, *19*(1). https://doi.org/10.1186/s12909-019-1550-x

Mulyono, D. (2021). The Influence of Learning Models and Learning Independence on Mathematics Learning Outcomes by Controlling Students' Early Abilities. *International Journal of Electronic Mathematics Education*, *12*(3). https://doi.org/10.29333/iejme/642

Nu'man, M., Retnawati, H., Sugiman, S., & Jailani, J. (2021). Measuring self-regulated learning in a STEM framework: A confirmatory factor analysis. *European Journal of Education Research*, *10*(4) 2067-2077. https://doi.org/10.12973/EU-JER.10.4.2067

Nurhayati, N., Jamaris, J., & Marsidin, S. (2022). Strengthening the Profile of Pancasila Students in the Independent Learning Curriculum in Elementary Schools. *International Journal of Education and Social Sciences and Humanities (IJHESS), 1*(6) 976-988. https://doi.org/10.55227/ijhess.v1i6.183

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulating learning in students. *Review of Educational Psychology*, *16*(4), 385-407. https://doi.org/10.1007/s10648-004-0006-x

Polit, D. F., & Beck, C. T. (2006). Content validity index: Are you sure you know what's being reported? Criticism and recommendations. *Research in Nursing and Health*, *29*(5), 489-497. https://doi.org/10.1002/nur.20147

Rawani, D., Indra Putri, R. I., Zulkardi, & Susanti, E. (2024). Application of the Indonesian realistic mathematics learning approach (PMRI) using the context of South Sumatera dance to improve numeracy literacy skills. AIP Conference Proceedings, 3046(1). https://doi.org/10.1063/5.0194767

Rusydiyah, E. F., Purwati, E., & Prabowo, A. (2020). How to use digital literacy as a learning resource for teacher candidates in Indonesia. Jurnal Cakrawala Pendidikan, 39(2), 305–318. https://doi.org/10.21831/cp.v39i2.30551

Schamberger, T., Schuberth, F., & Henseler, J. (2023). Confirmatory composite analysis in human development research. *International Journal of Behavioral Development*, *47*(1) 89-100. https://doi.org/10.1177/01650254221117506

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting of structural equation modeling results and confirmation factor analysis: A review. In *Journal of Educational Research*. 99(6), 323-338. https://doi.org/10.3200/JOER.99.6.323-338

Schubert, A. L., Hagemann, D., Voss, A., & Bergmann, K. (2017). Evaluation of the compatibility of diffusion models with mean square root of approach errors. *Journal of Mathematical Psychology*, *77, 29-45*. https://doi.org/10.1016/j.jmp.2016.08.004

Shi, D., Maydeu-Olivares, A., & DiStefano, C. (2018). Relationship between standard root square mean residue and incorrect model specification in factor analysis models. *Multivariate Behavior Research*, *53*(5), 676-694. https://doi.org/10.1080/00273171.2018.1476221

Sousa-Ferreira, R., Campanari-Xavier, R. A., & Rodrigues-Ancioto, A. S. (2021). Virtual reality as a tool for primary and vocational education. *General Scientific Journal of José María Cordova*, *19*(33). https://doi.org/10.21830/19006586.728

Sping, R. A. (2009). *Is there a way to increase the Average Variance Extracted (AVE) on the Latent Variable (LV) X (Revised)?. http://home.att.net/~rpingjr/ImprovAVE1.doc*

Surucu, L., & Maslakci, A. (2020). Validity and Reliability in Quantitative Research. *Business & Management Studies: International Journal*, *8*(3) 2694-2726. https://doi.org/10.15295/bmij.v8i3.1540

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association. https://doi.org/10.1037/10694-000

Valentini, F., & Damásio, B. F. (2016). Extracted Mean Variance and Composite Reliability: Reliability Coefficients. *Psychology: Theory and Research*, *32*(2). https://doi.org/10.1590/0102-3772e322225

Wahyuni, V. (2023). Aiken index analysis to determine the contents validity of project assessment instrument in learning Islamic religion and character in the independent curriculum. *Sustainable:*

Journal of Education Quality Studies, 6(2), 565-569. https://doi.org/10.32923/kjmp.v6i2.4043

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation. 13–39. https://doi.org/10.1016/B978-012109890-2/50031-7