Lensa: Jurnal Kependidikan Fisika
Penerbit: Universitas Pendidikan Mandalika
https://e-journal.undikma.ac.id/index.php/Lensa/index

June 2026, Vol. 14, No. 1
e-ISSN: 2686-0937
pp. 1–21

Construct Validation of a Critical Thinking Disposition Instrument for Physics Education Students: An Exploratory and Confirmatory Factor Analysis Approach

Ni Nyoman Sri Putu Verawati*, Joni Rokhmat, Ahmad Harjono, Muh Makhrus
Physics Education Department, Universitas Mataram, Mataram, Indonesia
**Corresponding Author:* veyra@unram.ac.id

Abstract

This study examined the construct validity of a critical thinking disposition instrument for Physics Education students using an Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) approach. The instrument measured four dimensions of critical thinking disposition, namely open-mindedness, truth-seeking, analyticity, and inquisitiveness, represented by 16 observed indicators. The study involved 250 undergraduate students enrolled in Physics Education programs from 10 public and private universities in Indonesia. Descriptive analysis showed that the mean score of each dimension was 3.00, with standard deviations ranging from 1.07 to 1.10 and Cronbach's alpha values between 0.745 and 0.784. The data met the requirements for factor analysis, as indicated by a Kaiser-Meyer-Olkin value of 0.863, a significant Bartlett's Test of Sphericity (χ^2 = 1247.060, $p < .001$), and a total reliability coefficient of 0.858. EFA identified a four-factor structure explaining 60.20% of the total variance. CFA further confirmed the adequacy of the measurement model, with all standardized factor loadings exceeding 0.70, composite reliability ranging from 0.847 to 0.857, and average variance extracted ranging from 0.580 to 0.600. These findings indicate that the instrument demonstrates satisfactory construct validity and reliability and is suitable for measuring critical thinking disposition among Physics Education students.

Keywords: Critical thinking disposition; Construct validity; Physics education students; Exploratory factor analysis; Confirmatory factor analysis

How to cite: Verawati, N. N. S. P., Rokhmat, J., Harjono, A., & Makhrus, M. (2026). Construct Validation of a Critical Thinking Disposition Instrument for Physics Education Students: An Exploratory and Confirmatory Factor Analysis Approach. *Lensa: Jurnal Kependidikan Fisika*, *14*(1), 1–21. https://doi.org/10.33394/j-lkf.v14i1.20097

INTRODUCTION

Critical thinking is one of the key competencies in higher education, particularly in science education, which requires the ability to analyze information, evaluate evidence, construct arguments, and make rational decisions. In physics learning, these demands become even more apparent because students are not only expected to master concepts but also to interpret phenomena, examine the consistency of explanations, and solve problems through scientific reasoning. However, critical thinking is not solely related to cognitive ability. A person may possess strong analytical skills but may not necessarily have the tendency to use those skills consistently when facing academic problems or situations that require intellectual judgment. Therefore, critical thinking disposition is an important aspect that deserves attention in higher education, including in the preparation of future physics teachers (Ali & Awan, 2021; Sireerat et al., 2022).

Critical thinking disposition refers to an individual's internal tendency to think openly, seek truth based on evidence, examine issues analytically, and demonstrate intellectual curiosity. Unlike critical thinking skills, which emphasize cognitive performance, disposition emphasizes a person's willingness and readiness to employ critical thinking processes in real situations. The literature suggests that critical thinking disposition is associated with the quality of reasoning, problem-solving, and decision-making across various fields. It should therefore not be viewed as a minor complement to critical thinking ability, but rather as a foundation that influences whether such ability is actually activated in learning and professional practice (Anders et al., 2019; Chen et al., 2020; Oh et al., 2025).

In the context of physics education, attention to critical thinking disposition becomes even more relevant because students in this field are prepared to become future teachers who are expected not only to master subject matter, but also to cultivate a culture of thinking among their learners. Physics Education students are expected to deal with scientific uncertainty, evaluate evidence, consider alternative explanations, and justify the reasoning behind their conclusions. Such tendencies are essential in physics learning because the field is closely related to inquiry, data-based argumentation, modeling, and problem-solving that do not always follow a single fixed procedure. Several studies have also shown that critical thinking disposition is associated with engagement in complex tasks, inquiry-based learning, and academic performance in science education and related contexts (Akbaş, 2021; Maison et al., 2019; Saputri et al., 2022).

Despite its importance, measuring critical thinking disposition is not a simple task. This construct is latent in nature, meaning that it cannot be observed directly but must be represented through observable indicators in the form of instrument items. This gives rise to a fundamental methodological issue. An instrument that appears conceptually sound or has undergone content validation does not necessarily possess an internal structure that accurately reflects the intended construct. For this reason, the development of a critical thinking disposition instrument requires sufficient evidence of construct validity so that the items are not only theoretically relevant but also function empirically as representations of the hypothesized latent construct (Leach, 2019; Moghaddam, 2019).

The need for an empirically validated instrument becomes even more urgent when the target of measurement is a specific population. Physics Education students have academic and professional characteristics that differ from those of school students or university students in other disciplines. They study in an environment that requires the integration of conceptual knowledge, scientific processes, and pedagogical orientation. Therefore, the use of a general instrument without adequate testing may produce measurements that are not sufficiently sensitive to the context of physics learning. Previous studies have also indicated that instrument development in a particular domain should consider the specificity of the context, the users, and the epistemic practices underlying it so that the resulting measurements are more meaningful and can be interpreted appropriately (Cui & Zhao, 2024; Emiliannur et al., 2017; Suprapto, 2019).

In the current study, the measurement of critical thinking disposition is focused on four main dimensions: open-mindedness, truth-seeking, analyticity, and inquisitiveness. These four dimensions were selected because they conceptually represent students' tendencies to consider alternative perspectives, seek evidence before accepting a claim, evaluate issues logically, and demonstrate intellectual interest in knowledge and inquiry. This structure aligns with the demands of physics learning, which requires openness to data, consistency in reasoning, analytical precision, and a willingness to continue exploring problems. Accordingly, this study is based on the assumption that measuring critical thinking disposition among Physics Education students must be grounded in a clear conceptual framework and examined through adequate psychometric procedures.

From a methodological perspective, studies on instrument development in education increasingly emphasize the importance of stepwise testing of the internal structure of measurement tools. Within this framework, Exploratory Factor Analysis (EFA) is useful for exploring latent patterns among items, whereas Confirmatory Factor Analysis (CFA) is used to test whether the hypothesized measurement model is supported empirically. The combination of these two approaches is considered important because it allows for stronger validation than relying solely on expert judgment or internal reliability. Various studies on instrument development have shown that the EFA-CFA approach provides a more solid basis for evaluating construct representation, factor structure stability, and instrument suitability within its intended context of use (Li et al., 2022; Luo & Zou, 2024; Xu et al., 2023).

Based on this background, the current study positions construct validation as an important step in developing a critical thinking disposition instrument for Physics Education students. A well-validated instrument is not only useful for psychometric research, but may also be employed for initial diagnosis, learning evaluation, curriculum development, and the assessment of educational interventions aimed at fostering productive intellectual attitudes among future physics teachers. By contrast, if the measurement is conducted using an instrument without a strong validity foundation, the interpretation of students' disposition profiles will remain weak and difficult to use as a basis for academic decision-making. Therefore, the development of an instrument that is both conceptually relevant and empirically valid constitutes a real need in physics education research (Baş & Bolat, 2022; Stewart & Dempsey, 2005).

Study Purpose and Research Questions

The current study aims to develop and validate the construct of a critical thinking disposition instrument for Physics Education students through an Exploratory Factor Analysis and Confirmatory Factor Analysis approach. More specifically, the study seeks to determine whether the developed items genuinely represent the four theoretical dimensions of open-mindedness, truth-seeking, analyticity, and inquisitiveness, and to examine whether the resulting measurement model demonstrates an acceptable factor structure, good construct reliability, and adequate convergent validity. This objective is consistent with current trends in educational instrument development that require stronger empirical evidence so that measurement tools can be used appropriately in both research and learning evaluation (Luo & Zou, 2024; Mirza et al., 2022). Based on this objective, the research questions of the study are as follows:

1. Does the developed critical thinking disposition instrument demonstrate adequate preliminary characteristics and meet the requirements for further factorial analysis among Physics Education students?
2. What factor structure emerges empirically from the critical thinking disposition instrument through Exploratory Factor Analysis, and to what extent does this structure correspond to the theoretical four-dimensional framework?

3. Does the measurement model obtain empirical support at the Confirmatory Factor Analysis stage, indicating that the instrument is valid and reliable for measuring critical thinking disposition among Physics Education students?

Novelty of the Study

The primary novelty of this study lies in its specific focus on developing and validating a critical thinking disposition instrument for Physics Education students. Many studies on critical thinking place greater emphasis on ability or cognitive performance, whereas the dispositional dimension often receives more limited attention. Yet the tendency to use reasoning in an open, evidence-based, and reflective manner is an important element in determining whether critical thinking ability is actually present in learning practice. By positioning Physics Education students as a clearly defined target population, this study offers a more contextualized focus than the use of general instruments that are not specifically designed for the needs of future physics teachers. Such an approach is important because the relevance of an instrument to the domain context and user characteristics strongly influences the meaning of the measurement outcomes (Galindo-Domínguez et al., 2023; Umar, 2023).

A further novelty lies in the validation approach employed. The current study does not stop at content validation, but extends the process to testing the internal structure of the instrument through a combination of EFA and CFA. This approach reflects a more systematic effort to build evidence of construct validity, because the relationships between items and latent constructs are not merely assumed on theoretical grounds, but are also examined through exploratory and confirmatory factor analyses. In the context of educational instrument development, such a procedure is important for ensuring that the instrument has a stronger psychometric foundation, can be interpreted conceptually, and is appropriate for use with a specific target population. Thus, the novelty of this study lies not only in its focused population, but also in its commitment to a more comprehensive construct validation procedure in the field of physics education (Li et al., 2022; Manggaberani & Putro, 2024).

METHODS

This study employed a quantitative approach with an instrument development and validation design. The main objective was to examine the construct validity of an instrument intended to measure the critical thinking disposition of Physics Education students through two stages of factor analysis, namely *Exploratory Factor Analysis* (EFA) and *Confirmatory Factor Analysis* (CFA). This design was selected because an instrument used to measure a latent psychological construct should not rely solely on content validity, but also needs empirical evidence showing that the indicators genuinely represent the construct being measured.

Research Design

The study was conducted as an empirical instrument validation study. The procedure included defining the theoretical construct, developing the instrument blueprint, writing the item statements, conducting expert judgment for content validation, administering the instrument to respondents, and analyzing the data using EFA and CFA. The main focus of the study was to test the internal structure of the instrument and to confirm whether the theoretical dimensions of critical thinking disposition were supported by empirical data.

This approach was considered appropriate because critical thinking disposition is a latent construct that cannot be observed directly. Therefore, it must be measured through a set of observed indicators reflecting students' tendencies to be open-minded, seek evidence, analyze problems carefully, and demonstrate intellectual curiosity. In this context, factor analysis provides an appropriate statistical framework for evaluating the relationship between observed items and their latent dimensions.

Participants

The participants in this study were 250 undergraduate students enrolled in Physics Education programs from 10 public and private universities in Indonesia. This sample size was considered adequate for factor analysis because it satisfies commonly accepted recommendations for testing measurement models with a moderate number of observed variables. With 16 items included in the instrument, this sample provided sufficient information to estimate inter-item relationships, identify the latent factor structure, and assess the stability of the measurement model. The participants represented the target population for whom the instrument was intended, namely preservice physics teachers in higher education. Data obtained from these respondents were used for descriptive analysis, factorability testing, EFA, and CFA.

Construct and Operational Definition

The instrument was developed based on four dimensions of critical thinking disposition, namely open-mindedness, truth-seeking, analyticity, and inquisitiveness. These four dimensions were selected because, conceptually, they represent students' internal tendencies to consider alternative viewpoints, seek evidence before accepting a claim, examine problems logically, and show intellectual interest in knowledge and inquiry.

Operationally, critical thinking disposition was defined as students' tendency to demonstrate openness, evidence orientation, analytical reasoning, and intellectual curiosity when dealing with information, problems, or decision-making situations. This construct was measured using 16 observed indicators distributed across the four dimensions, with each dimension represented by four items. The open-mindedness dimension refers to students' willingness to consider different viewpoints and avoid rejecting ideas prematurely. The truth-seeking dimension reflects students' tendency to search for facts and evidence before accepting a statement. The analyticity dimension represents students' tendency to examine relationships, evaluate arguments, and process information logically. The inquisitiveness dimension refers to students' interest in exploring ideas, asking questions, and understanding issues more deeply.

Instrument Development

The research instrument was developed in the form of a closed-ended questionnaire using a five-point Likert scale: 1 = strongly disagree, 2 = disagree, 3 = undecided, 4 = agree, and 5 = strongly agree. This response format was selected because it is suitable for measuring dispositions, attitudes, and psychological tendencies. Each statement was designed to represent one indicator relevant to the intended latent dimension.

At the initial stage, an instrument blueprint was developed based on the operational definition of each dimension. From this blueprint, 16 items were generated. The initial composition of the instrument consisted of OP1 to OP4 for open-mindedness, TR1 to TR4 for truth-seeking, AN1 to AN4 for analyticity, and IN1 to IN4 for inquisitiveness.

Before empirical testing, the items were reviewed by experts to assess their relevance, clarity, and appropriateness for the target respondents. This content validation process was intended to ensure that the wording of each item genuinely reflected the construct being measured and could be clearly understood by university students. Feedback from the experts was used to revise the wording of the items and strengthen the conceptual alignment between indicators and dimensions.

Data Collection Procedure

After the instrument had been reviewed and refined, the questionnaire was administered to the participants. Before completing the questionnaire, respondents were informed about the purpose of the study, the confidentiality of their responses, and the procedures for completing the instrument. They were also told that there were no right or wrong answers, so they were encouraged to respond honestly according to their own tendencies and experiences.

The completed responses were then organized into a data matrix consisting of 250 cases and 16 items. These data were subsequently analyzed to examine descriptive characteristics, test the suitability of the data for factor analysis, identify the factor structure exploratorily, and confirm the measurement model of students' critical thinking disposition.

Data Analysis

Data analysis was conducted in four stages: descriptive statistics, factorability testing, Exploratory Factor Analysis (EFA), and Confirmatory Factor Analysis (CFA). The first stage involved descriptive analysis to examine the general characteristics of responses for each dimension. The reported statistics included the mean, standard deviation, and Cronbach's alpha coefficient. These statistics were used to evaluate central tendency, score variability, and the internal consistency of the instrument before conducting factor analysis.

Before performing EFA, the suitability of the data for factor analysis was tested using the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) and Bartlett's Test of Sphericity. The KMO index was used to determine whether the correlation matrix was sufficiently compact to produce reliable factors, while Bartlett's test was used to assess whether the correlation matrix significantly differed from an identity matrix. The data were considered suitable for factor analysis when the KMO value exceeded 0.50 and Bartlett's test was significant at $p < 0.05$. In addition, the overall reliability of the instrument was examined using Cronbach's alpha.

EFA was conducted to identify the empirical factor structure of the 16-item instrument. The analysis was based on the inter-item correlation matrix, with factor extraction performed using a principal factor extraction method. To obtain a simpler and more interpretable structure, orthogonal varimax rotation was applied. The number of retained factors was determined based on three considerations:

eigenvalues greater than 1, visual inspection of the scree plot, and the interpretability of the factor structure in relation to the theoretical model. An item was considered to load adequately on a factor when its primary loading was at least 0.50 and it did not show problematic cross-loading on other factors. Items with low loadings or relatively high cross-loadings were reviewed further before being retained in the model.

The EFA results were interpreted through eigenvalues, cumulative explained variance, the *scree plot*, the dominant factor structure, and the loading pattern of each item. These results provided the empirical basis for evaluating whether the hypothesized four-dimensional structure was supported at the exploratory level.

Confirmatory Factor Analysis

After the exploratory structure had been identified, CFA was conducted to confirm the measurement model. CFA specified a first-order model with four factors, in which each latent construct was represented by four observed indicators. The measurement model was reflective, meaning that the latent constructs were assumed to explain the observed responses to their indicators. The measurement model can be expressed as follows:

$$x_i = \lambda_i \xi + \delta_i \qquad (1)$$

where x_i represents an observed indicator, λ_i is the factor loading, ξ is the latent construct, and δ_i is the measurement error. In this specification, each indicator loaded on only one latent factor, while correlations among latent factors were allowed.

CFA evaluation focused on *standardized factor loadings, Composite Reliability* (CR), and *Average Variance Extracted* (AVE). Standardized loadings were used to assess the contribution of each indicator to its latent construct. Indicators were considered adequate when they had loading values of at least 0.50, with values above 0.70 considered strong. CR values above 0.70 indicated good construct reliability, while AVE values above 0.50 indicated adequate convergent validity.

Decision Criteria

Decisions regarding construct validity were based on several criteria. First, the data had to meet the assumptions for factor analysis through an adequate KMO value and a significant Bartlett's test. Second, the EFA results had to show a factor structure that was interpretable and consistent with the theoretical framework. Third, the CFA results had to demonstrate that each item loaded adequately on its intended latent construct. Fourth, each construct had to meet the minimum standards for CR and AVE. Items showing weaker loading patterns or tendencies toward cross-loading could still be retained if they remained statistically acceptable and theoretically meaningful within the confirmatory model.

Analytical Software

Data analysis was conducted using statistical software. Descriptive statistics, KMO, Bartlett's test, and EFA were performed using standard statistical analysis software. CFA and the visualization of the measurement model were conducted using structural equation modeling software. These tools enabled estimation of the factor structure, evaluation of item performance, and assessment of construct reliability and convergent validity.

Ethical Considerations

This study was conducted in accordance with general ethical principles in educational research. Participants were informed about the purpose of the study, the confidentiality of their responses, and the use of the data solely for academic purposes. Participation was voluntary, and respondents' identities were kept confidential throughout the research process.

RESULTS AND DISCUSSION

Results

To provide an initial overview of the data characteristics prior to testing the construct structure, Table 1 presents the descriptive statistics for each dimension of Physics Education students' critical thinking disposition, including the mean, standard deviation, and internal reliability coefficient.

Table 1. Descriptive statistics and reliability of each dimension of CT disposition

Dimension	Mean	SD	Cronbach's Alpha
Open-mindedness	3.00	1.10	0.784
Truth-seeking	3.00	1.07	0.745
Analyticity	3.00	1.08	0.757
Inquisitiveness	3.00	1.09	0.765

Table 1 shows that the mean scores of all dimensions were 3.00. This pattern indicates that Physics Education students' responses tended to fall within the moderate category, suggesting that the data were not concentrated at either the very low or very high response categories. Thus, the score distribution can be considered sufficiently balanced to represent variation in Physics Education students' critical thinking disposition across the four measured dimensions.

The standard deviations, ranging from 1.07 to 1.10, indicate that respondents' answers showed adequate variability. In terms of reliability, all dimensions had Cronbach's Alpha values above 0.70, indicating that the items within each dimension demonstrated good internal consistency. The open-mindedness dimension had the highest reliability, whereas truth-seeking had the lowest, although it still remained within the acceptable range. These findings provide an initial basis for proceeding to factor analysis.

Before conducting factor analysis, it was necessary to ensure that the inter-item correlation matrix met the assumptions required for further analysis. Table 2 presents the results of the factorability tests using the Kaiser-Meyer-Olkin (KMO) measure, Bartlett's Test of Sphericity, and the total reliability of the instrument.

Table 2. Results of factorability testing

Factorability Indicator	Value
Number of respondents	250
Number of items	16
KMO	0.863
Bartlett's Test χ^2	1247.060
df	120
Sig. Bartlett	0.000
Total Cronbach's Alpha	0.858

Table 2 shows that the KMO value was 0.863. This value indicates that the data had good sampling adequacy for factor analysis. The higher the KMO value, the better the inter-item correlation pattern for forming a stable factor structure. Therefore, the KMO value in this study confirms that the data were adequate for exploratory analysis.

Bartlett's Test of Sphericity yielded a χ^2 value of 1247.060 with a significance level of 0.000. This finding indicates that the correlation matrix significantly differed from an identity matrix, meaning that the items were sufficiently correlated to form factors. The total instrument reliability of 0.858 also indicates that the instrument as a whole had good internal consistency. Based on these results, the data were considered suitable for proceeding to the Exploratory Factor Analysis stage.

To determine the number of factors to retain, Figure 1 presents the scree plot from the Exploratory Factor Analysis based on the distribution of eigenvalues across components.

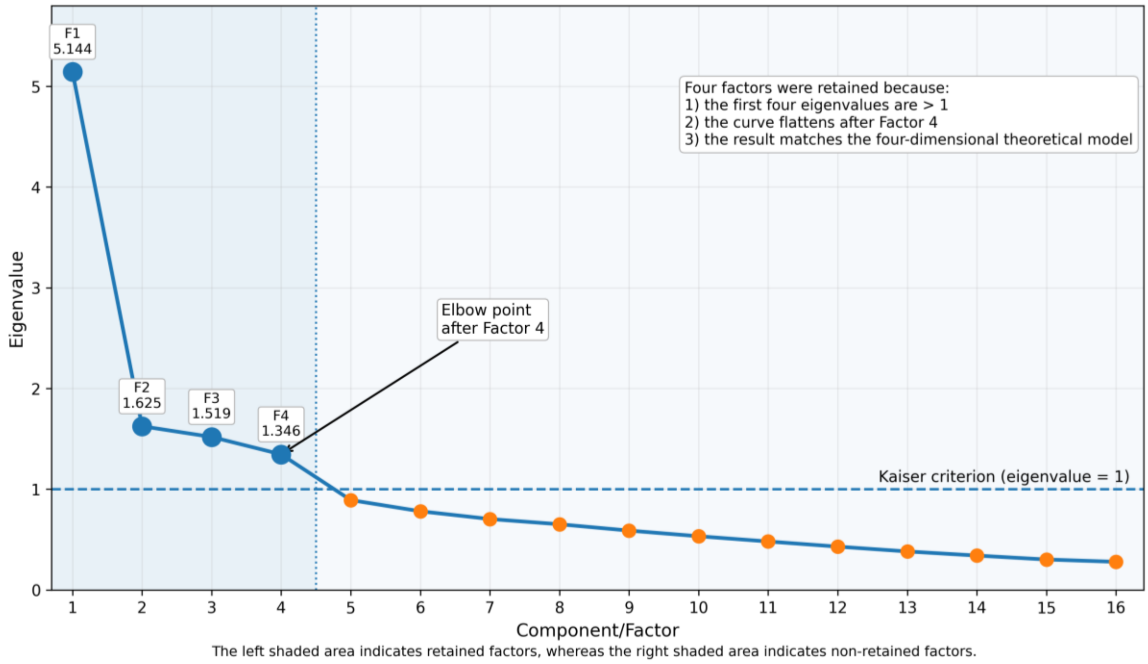

Figure 1. Scree plot of the Exploratory Factor Analysis (EFA)

Figure 1 shows that the eigenvalues declined sharply across the first four components, after which the curve began to flatten. This pattern indicates an elbow point after the fourth factor, suggesting that the first four factors were the most meaningful in explaining data variance. Beyond the fourth factor, additional components no longer made a substantial contribution to the construct structure.

In addition to the visual pattern of the curve, the decision to retain four factors was also supported by the Kaiser criterion, as only the first four factors had eigenvalues greater than 1. Thus, the scree plot served not only as a graphical illustration but also as empirical justification for the number of retained factors. This finding is consistent with the theoretical framework underlying the instrument development, namely that Physics Education students' critical thinking disposition is represented by four major dimensions: open-mindedness, truth-seeking, analyticity, and inquisitiveness.

To clarify the basis for factor retention, Table 3 presents the eigenvalues, the proportion of variance explained by each factor, and the cumulative percentage of explained variance.

Table 3. Eigenvalues and proportion of explained variance

Factor	Eigenvalue	Variance Explained (%)	Cumulative (%)
Factor 1	5.144	32.15	32.15
Factor 2	1.625	10.15	42.30
Factor 3	1.519	9.49	51.79
Factor 4	1.346	8.41	60.20

Table 3 shows that the first four factors had eigenvalues greater than 1.00, thus meeting the Kaiser criterion for retention. The first factor contributed the largest proportion of total variance, namely 32.15%, while the second, third, and fourth factors contributed 10.15%, 9.49%, and 8.41%, respectively. This pattern suggests that the resulting factor structure was not based on a single dominant dimension, but rather supported by several relatively balanced constructs.

Cumulatively, the four factors explained 60.20% of the total variance in the data. This indicates that the four-factor model had reasonably good explanatory power for variations in Physics Education students' responses. In the context of psychometric instrument development, a cumulative variance percentage above 50% is generally considered adequate to support the existence of a strong and conceptually interpretable factor structure.

After the number of factors had been determined, the EFA structure was further visualized through a path diagram to illustrate the relationship between the latent factors and the items with the strongest loadings. Figure 2 presents the EFA path diagram based on the dominant factors and their primary loading values.

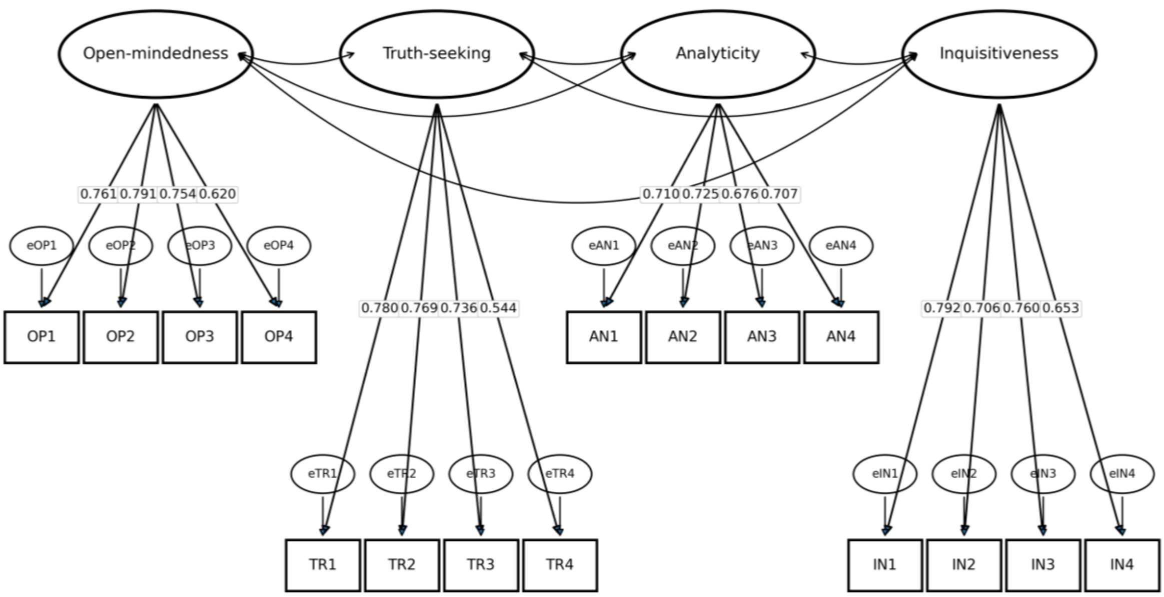

Figure 2. Path diagram of the Exploratory Factor Analysis (EFA)

Figure 2 shows that the items within each factor were fairly clearly concentrated on the constructs that corresponded to the instrument's theoretical framework. The open-mindedness factor was represented by OP1, OP2, OP3, and OP4; the truth-seeking factor by TR1, TR2, TR3, and TR4; the analyticity factor by AN1, AN2, AN3, and AN4; and the inquisitiveness factor by IN1, IN2, IN3, and IN4. This diagram indicates that the hypothesized four-factor structure received empirical support at the exploratory stage.

The loading values displayed on each path also indicate that most items had sufficiently strong associations with their dominant factors. Items in the open-mindedness and truth-seeking dimensions showed slightly more variation than those in the other two dimensions, particularly OP4 and TR4, which had lower loadings than other items within their respective groups. Nevertheless, overall, the diagram confirms that the separation among constructs had been reasonably well established at the EFA stage.

To present the EFA results in more detail and quantitatively, Table 4 displays the primary loading of each item on its dominant factor, the highest cross-loading, and the preliminary decision regarding the construct validity of each item.

Table 4. EFA results for item construct validity

Item	Dominant Factor	Primary Loading	Highest Cross-loading	Decision
OP1	Open-mindedness	0.761	0.214	Valid
OP2	Open-mindedness	0.791	0.201	Valid
OP3	Open-mindedness	0.754	0.226	Valid
OP4	Open-mindedness	0.620	0.421	Review
TR1	Truth-seeking	0.780	0.236	Valid
TR2	Truth-seeking	0.769	0.228	Valid
TR3	Truth-seeking	0.736	0.241	Valid
TR4	Truth-seeking	0.544	0.402	Review
AN1	Analyticity	0.710	0.278	Valid
AN2	Analyticity	0.725	0.244	Valid
AN3	Analyticity	0.676	0.291	Valid
AN4	Analyticity	0.707	0.233	Valid
IN1	Inquisitiveness	0.792	0.210	Valid
IN2	Inquisitiveness	0.706	0.268	Valid
IN3	Inquisitiveness	0.760	0.219	Valid
IN4	Inquisitiveness	0.653	0.298	Valid

Table 4 indicates that the majority of items had primary loadings above 0.60 on the intended factor. This suggests that most indicators had strong relationships with the constructs they were intended to represent. OP1, OP2, OP3, TR1, TR2, TR3, AN1, AN2, AN4, IN1, and IN3 were among the items with relatively strong loadings, while AN3, IN2, and IN4 remained within an adequate range to be retained.

Two items, namely OP4 and TR4, required particular attention because, in addition to having lower primary loadings than other items within their dimensions, they also showed relatively high cross-loadings. This condition suggests possible conceptual overlap with other factors. However, because their primary loadings remained above the minimum acceptable threshold, these items were not immediately eliminated and were retained for further testing at the confirmatory stage. Overall, the EFA results showed that 14 of the 16 items met the initial construct validity criteria very well, while two items required further review.

After the four-factor structure had been identified at the exploratory stage, the analysis proceeded to the confirmatory stage to assess whether the measurement model was truly supported by the data. Figure 3 presents the Confirmatory Factor Analysis (CFA) path diagram used to confirm the relationships between the latent constructs and their indicators.

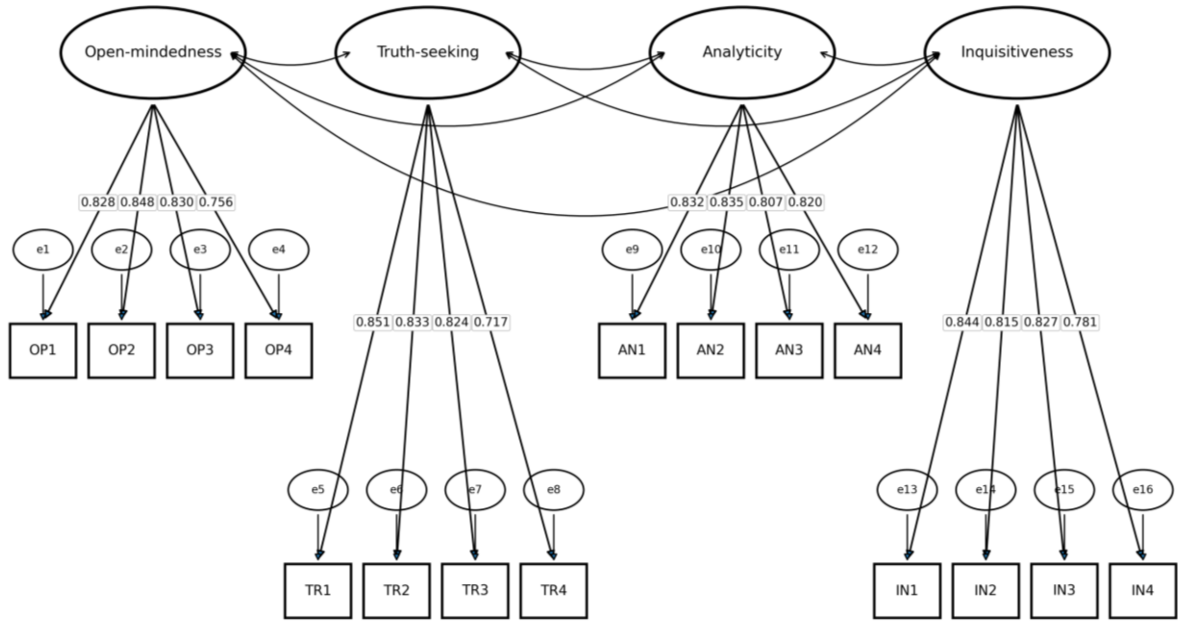

Figure 3. Confirmatory Factor Analysis (CFA) path diagram of Physics Education students' critical thinking disposition

Figure 3 presents a first-order measurement model consisting of four latent constructs and sixteen observed indicators. Each latent construct was directly linked to four indicators representing it. In addition, measurement residuals for each item were included in the model, making the model structure complete. The bidirectional arrows among latent constructs indicate that the dimensions of critical thinking disposition were treated as distinct yet correlated constructs.

Visually, the CFA model shows a clear specification consistent with the EFA results. Each indicator was assigned to only one latent construct, and there were no cross-paths from indicators to other constructs. Thus, the model tested at the confirmatory stage genuinely represented the four-factor structure identified during the exploratory stage. This diagram strengthens the argument that the instrument was structured as a stable and well-defined reflective model.

To assess the strength of the relationships between indicators and latent constructs in the confirmatory model, Table 5 presents the standardized loading values from the CFA for all items.

Table 5. Standardized loadings from the CFA

Construct	Item	Standardized Loading	Decision
Open-mindedness	OP1	0.788	Very good
Open-mindedness	OP2	0.823	Very good
Open-mindedness	OP3	0.766	Very good
Open-mindedness	OP4	0.739	Very good
Truth-seeking	TR1	0.805	Very good
Truth-seeking	TR2	0.761	Very good
Truth-seeking	TR3	0.741	Very good
Truth-seeking	TR4	0.704	Very good
Analyticity	AN1	0.800	Very good
Analyticity	AN2	0.758	Very good
Analyticity	AN3	0.757	Very good

Construct	Item	Standardized Loading	Decision
Analyticity	AN4	0.727	Very good
Inquisitiveness	IN1	0.784	Very good
Inquisitiveness	IN2	0.738	Very good
Inquisitiveness	IN3	0.785	Very good
Inquisitiveness	IN4	0.756	Very good

Table 5 shows that all indicators had standardized loadings above 0.70. This result indicates that all items had strong relationships with the latent constructs they measured. No item fell below the minimum acceptable threshold; therefore, all indicators could be retained in the final model. The highest loadings were found for OP2 and TR1, whereas the lowest loading was found for TR4, although it still remained in the very good category.

Compared with the EFA results, the CFA stage provided stronger confirmation of the instrument's stability. OP4 and TR4, which required review at the exploratory stage, still demonstrated adequate loadings at the confirmatory stage, namely 0.739 and 0.704, respectively. This finding suggests that although these two items initially indicated cross-loading tendencies during the exploratory stage, they were still able to represent their latent constructs adequately overall. Therefore, all 16 items were retained in the final measurement model.

In addition to evaluating indicator loadings, the CFA was also used to assess construct reliability and convergent validity. Table 6 presents the values of Composite Reliability (CR) and Average Variance Extracted (AVE) for each latent construct.

Table 6. Values of CR and AVE

Construct	CR	AVE
Open-mindedness	0.857	0.600
Truth-seeking	0.849	0.584
Analyticity	0.847	0.580
Inquisitiveness	0.854	0.595

Table 6 shows that all constructs had CR values above 0.70. This indicates that the indicators within each dimension had good consistency in reflecting their latent constructs. The relatively balanced CR values across constructs also suggest that measurement quality was distributed evenly rather than concentrated in only one dimension. The AVE values for all four constructs were also above 0.50. This finding indicates that more than half of the variance of the indicators could be explained by their respective latent constructs, meaning that the model's convergent validity was achieved. Thus, the measurement model of Physics Education students' critical thinking disposition was not only reliable but also demonstrated adequate construct quality for use in subsequent research.

To summarize the entire sequence of instrument testing results from the descriptive stage to the confirmatory stage, Table 7 presents the final synthesis of the quality of the critical thinking disposition instrument.

Table 7. Final summary of instrument quality for critical thinking disposition

Evaluation Aspect	Main Result	Decision
Internal reliability of dimensions	Alpha 0.745–0.784	Adequate
Total instrument reliability	Alpha = 0.858	Good

Evaluation Aspect	Main Result	Decision
Factorability	KMO = 0.863; Bartlett sig. = 0.000	Suitable
EFA factor structure	4 factors; cumulative var. = 60.20%	Supported
EFA item validity	14 items valid; 2 items reviewed	Mostly valid
CFA loadings	All items > 0.70	Very good
CFA construct reliability	CR 0.847–0.857	Good
CFA convergent validity	AVE 0.580–0.600	Adequate

Table 7 shows that the Physics Education students' critical thinking disposition instrument had good psychometric quality overall. At the initial stage, the data met all requirements for factor analysis. The EFA successfully identified four factors consistent with the theoretical model, with adequate cumulative variance. Although two items required attention at the exploratory stage, the majority of items were found to be valid and formed a clear factor structure.

At the confirmatory stage, all indicators showed strong loadings, good construct reliability, and adequate convergent validity. Therefore, the developed measurement model of critical thinking disposition can be considered appropriate for use. These results indicate that the combination of EFA and CFA provided a strong empirical basis for concluding that the instrument was able to measure the four dimensions of Physics Education students' critical thinking disposition consistently and validly.

Discussion

The findings of this study indicate that the critical thinking disposition instrument developed for Physics Education students has a clear construct structure and adequate psychometric support. In general, these findings reinforce the view that critical thinking disposition is not a simple, single attribute, but rather a multidimensional construct consisting of interconnected yet distinguishable tendencies of thought. In the current study, the four proposed dimensions, namely open-mindedness, truth-seeking, analyticity, and inquisitiveness, received support at both the exploratory and confirmatory stages. This is important because it suggests that the thinking orientation of Physics Education students can be mapped into components that are conceptually relevant to physics learning, including openness to evidence, analytical precision, willingness to question claims, and intellectual curiosity. Accordingly, the results of this study are in line with the literature that positions critical thinking disposition as a construct broader than mere cognitive performance and one that should be assessed through an empirically validated internal structure (Roberts et al., 2015; Bakhtiari-Dovvombaygi et al., 2024).

The initial quality of the instrument in this study also demonstrated a sufficient basis for further psychometric testing. The descriptive statistics and initial reliability suggested that students' responses were not concentrated at extreme points, allowing the items to differentiate levels of disposition across respondents. From a measurement perspective, this condition is essential because a disposition instrument should not merely identify the presence of certain attitudes, but should also be sensitive to different degrees of tendency. This finding is consistent with Anders et al. (2019), who emphasized that a sound critical thinking instrument should demonstrate adequate discriminative capacity to be meaningfully applied in educational contexts. In a more applied setting, Chen et al. (2020) also showed that critical thinking disposition is associated with professional competence and the quality of decision-making, which means that the instrument used to assess it must be sufficiently sensitive to individual variation. Viewed from this perspective, the relatively balanced response distribution in the current study serves as an early indication that the instrument can capture the spectrum of critical thinking disposition among Physics Education students.

The results of the factorability test further strengthened this argument. Data that met the requirements for factorial analysis indicate that the developed items did not stand independently, but rather formed a systematic pattern of relationships. Methodologically, this supports the idea that construct validation cannot be reduced to content validation alone. Moghaddam (2019), in developing an instrument to measure teachers' perceptions, argued that the conceptual quality of items must be accompanied by internal evidence regarding the relationships among indicators. A similar position was reported by Mirza et al. (2022), who demonstrated that factor analysis and internal consistency are essential foundations for ensuring that an instrument actually functions in accordance with its theoretical construction. Therefore, the factorability results in the current study should not be seen merely as a technical step, but also as evidence that the four proposed dimensions indeed had an empirical basis for further testing.

At the Exploratory Factor Analysis stage, this study identified four factors that were consistent with the initial theoretical framework. This finding is noteworthy because it differs from some previous studies that obtained two- or three-factor solutions in instruments measuring critical thinking disposition or related affective constructs. For example, Bakhtiari-Dovvombaygi et al. (2024) and Tep et al. (2021) showed that in some cultural and disciplinary contexts, the resulting structure tended to be more condensed. In several cross-context validation studies, simpler models sometimes produced a better statistical fit than more detailed ones. However, the current study showed that in the context of Physics Education students, the four-dimensional structure could still be retained. This difference should not be interpreted as an inconsistency that weakens the findings, but rather as an indication that the structure of critical thinking disposition may vary according to disciplinary context, respondent characteristics, and measurement goals. In other words, the results of this study support the argument that disposition constructs do not always need to be reduced to simpler models when more specific dimensions remain theoretically and empirically distinguishable.

The appropriateness of the four-factor structure in this study also has substantive significance for physics education. Physics learning demands more than the ability to solve problems. Students must be willing to accept the possibility that their initial explanations may be incorrect, be motivated to seek evidence, be able to analyze relationships among variables, and possess curiosity toward phenomena they do not yet understand. Therefore, the four-dimensional structure retained in this study may in fact be more representative of the domain of physics education than an overly general model. This interpretation is consistent with the view that instruments should be sensitive to domain and epistemic practice (Cui & Zhao, 2024; Emiliannur et al., 2017). Suprapto (2019) also showed that in physics education, students' learning perceptions and orientations are often shaped by the distinctive demands of the discipline, meaning that instruments developed contextually are likely to be more meaningful than general instruments applied across domains.

At the item level, the findings of this study showed that most indicators had strong loadings on their dominant factors. Nevertheless, the presence of two items that showed cross-loading tendencies at the EFA stage provides room for a more critical interpretation. In instrument development, such a situation does not always mean that the items are fatally problematic. Leach (2019), through a multi-analytic approach to critical thinking tests, showed that the factor structure of an instrument may display complex nuances that cannot always be captured adequately through a single criterion for item elimination. Similarly, Coelho et al. (2018) emphasized that decisions to retain or remove items should balance statistical parsimony and conceptual representation. In the current study, the decision to retain the sensitive items during the EFA stage proved justifiable because at the CFA stage these items still showed adequate contributions. This suggests that item evaluation in instrument development should indeed be conducted gradually, cautiously, and not in an overly mechanistic manner.

The Confirmatory Factor Analysis stage provided important reinforcement for the exploratory results. All indicators showed adequate contributions to their respective latent constructs, which means that the four-factor model did not merely emerge from an initial data pattern, but also received support when tested in a stricter measurement model. This finding is consistent with general recommendations in the instrument validation literature that EFA and CFA should be understood as complementary processes rather than interchangeable ones (Chen & Fan, 2022; Luo & Zou, 2024; Xu et al., 2023). Compared with validation that stops at EFA alone, the use of CFA in the current study clearly strengthens the evidence of construct validity. In this regard, the study makes an important contribution to instrument development in physics education, as it demonstrates that a two-stage approach can be applied productively to examine the representation of critical thinking disposition constructs.

Even so, from a methodological standpoint, one issue still deserves critical discussion. The modern instrument validation literature generally recommends the use of two independent samples, one for EFA and another for CFA, so that the confirmatory results are not excessively influenced by the same data pattern. This point is emphasized by Bakhtiari-Dovvombaygi et al. (2024), Jalali et al. (2024b), and Chen and Fan (2022), who showed that a two-sample design can reduce circularity and improve the strength of inference. The current study has not yet reached that level of rigor. Therefore, its findings are better interpreted as strong evidence of construct validity at an early stage of instrument development rather than as a fully independent final validation. This criticism is important not because it weakens the results, but because it places them proportionally. In other words, the instrument is already promising, but it still needs to be tested again with other samples to ensure the stability of the model.

In addition, recent literature also emphasizes the importance of testing measurement invariance when an instrument is intended for use across different groups. Studies such as Trigueros et al. (2020), Dávila-Villavicencio et al. (2024), and Papagiannopoulou et al. (2025) show that configural, metric, and even scalar invariance become important if the results of an instrument are to be compared across gender, institutions, or cultural backgrounds. The current study has not yet reached that stage. This means that the instrument has shown good internal quality, but its functional equivalence across subgroups has not yet been established. In the highly diverse context of Indonesian higher education, this limitation is important to acknowledge. Physics Education students from different universities may experience different academic environments, curricula, and learning cultures, so further validation through measurement invariance testing would greatly strengthen the utility of this instrument.

The aspects of reliability and convergent validity in this study also demonstrated strong results, and this is consistent with trends in other instrument validation studies. Dávila-Villavicencio et al. (2024), Haeruddin et al. (2020), and Jalali et al. (2024a) showed that a strong instrument is usually supported by a combination of adequate internal reliability and good indicators of construct convergence. In the current study, such support suggests that each dimension was not only internally consistent but also sufficiently focused on its own latent construct. This is important because a critical thinking disposition instrument would lose its meaning if its items were too diffuse or excessively overlapping.

Substantively, the findings of this study support the view that critical thinking disposition should be treated as an important part of the academic and professional readiness of future physics teachers. Maison et al. (2019) showed that the learning environment and self-regulation are closely related to how students learn physics. Akbaş (2021) also demonstrated that argumentation-based learning can strengthen students' critical thinking disposition. When these findings are read together with the results of the current study, it becomes clear that a validated instrument can serve as a strategic tool for mapping students' dispositional readiness before and after instructional interventions. In other words, the value of this instrument lies not only in its psychometric validity, but also in its potential for evaluating curricula, teaching strategies, and the development of future physics teachers who are more reflective and evidence-based.

Finally, this study also highlights the importance of domain-specific instruments in educational research. As noted by Lins-Kusterer et al. (2024) and Yuan et al. (2018), the validation of a scale should not rely solely on the translation or adoption of a general concept; attention must also be paid to the context of use, respondent characteristics, and the structure of meaning of the construct in the new setting. Within this framework, the current study makes a relevant contribution because it does not merely adapt a general idea of critical thinking disposition, but tests it in the specific context of Physics Education students. Thus, the main strength of this study lies in two points: first, its success in demonstrating that the instrument has a convincing construct foundation; second, its willingness to position critical thinking disposition as an important target of measurement in physics education.

CONCLUSION

This study successfully developed and validated the construct of a critical thinking disposition instrument for Physics Education students through an Exploratory Factor Analysis and Confirmatory Factor Analysis approach. Overall, the findings indicate that the developed instrument has good psychometric quality and a construct structure consistent with the proposed theoretical framework. The four main dimensions, namely open-mindedness, truth-seeking, analyticity, and inquisitiveness, received empirical support from the data at both the exploratory and confirmatory stages. At the preliminary stage, the instrument demonstrated adequate descriptive characteristics, good internal reliability across dimensions, and strong data suitability for factor analysis. The EFA results showed that the four-factor structure could be retained and had sufficient explanatory power for variations in students' responses, while the CFA results confirmed that all indicators made adequate contributions to their respective latent constructs. Therefore, this study provides evidence that the critical thinking disposition of Physics Education students can be measured through a clear, structured, and empirically supported four-dimensional model.

These findings are important both methodologically and substantively. From a methodological perspective, this study confirms that the development of educational instruments, particularly those intended to measure latent constructs such as critical thinking disposition, requires stepwise validation that goes beyond content validation to include both exploratory and confirmatory examinations of internal structure. From a substantive perspective, the findings show that measuring critical thinking disposition among Physics Education students should take into account the specific characteristics of physics education, which demands openness to evidence, logical analysis, truth-seeking, and intellectual curiosity. Accordingly, the instrument developed in this study is relevant not only for psychometric purposes, but also for learning evaluation, initial diagnosis of student characteristics, curriculum development, and the assessment of interventions aimed at strengthening a culture of critical thinking among future physics teachers. Nevertheless, the findings should still be interpreted within the limits of the study, particularly because the validation was conducted within a single respondent group and did not yet include cross-sample testing or invariance analysis across subgroups. Therefore, this instrument may be regarded as a measurement tool with a strong validity foundation at the early stage of development and one that is worthy of further refinement in future studies.

RECOMMENDATION

Based on the findings of this study, it is recommended that future research retest the instrument using broader and more diverse samples, including Physics Education students from different institutional contexts, so that the stability of the model can be examined more rigorously. Further studies should also consider using separate samples for EFA and CFA, conducting measurement invariance testing across groups, and linking critical thinking disposition scores with indicators of academic performance or more direct manifestations of critical thinking behavior. In addition, lecturers and program administrators may use this instrument as a basis for mapping students' critical thinking disposition profiles and for designing more responsive instructional strategies to foster open-mindedness, evidence orientation, analytical reasoning, and intellectual curiosity.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to all Physics Education students from various universities in Indonesia who participated in this study, as well as to the experts and all parties who provided valuable input during the instrument development process. Appreciation is also extended to the institutions and academic environment that supported the implementation of this study.

REFERENCES

Ali, G., & Awan, R. (2021). Thinking based instructional practices and academic achievement of undergraduate science students: Exploring the role of critical thinking skills and dispositions. *Journal of Innovative Sciences, 7*(1), 56–70. https://doi.org/10.17582/journal.jis/2021/7.1.56.70

Akbaş, Y. (2021). The effects of argumentation-based teaching approach on students' critical thinking disposition and argumentation skills: "Population in our country unit". *International Journal of Psychology and Educational Studies, 8*(1), 51–74. https://doi.org/10.17220/ijpes.2021.8.1.195

Anders, P., Stellrecht, E., Davis, E., & McCall, W. (2019). A systematic review of critical thinking instruments for use in dental education. *Journal of Dental Education, 83*(4), 381–397. https://doi.org/10.21815/jde.019.043

Bakhtiari-Dovvombaygi, H., Pourhasan, K., Rahmaty, Z., Zare-Kaseb, A., Abbaszadeh, A., Rashtbarzadeh, A., & Borhani, F. (2024). Evaluation of cross-cultural adaptation and validation of the Persian version of the critical thinking disposition scale: Methodological study. *BMC Nursing, 23*(463). https://doi.org/10.1186/s12912-024-02129-y

Baş, M., & Bolat, Y. (2022). The impact of cognitive competence on critical thinking skills: An educational science study with school counsellors. *Education Quarterly Reviews, 5*(4), 53–64. https://doi.org/10.31014/aior.1993.05.04.605

Chen, Q., Liu, D., Zhou, C., & Tang, S. (2020). Relationship between critical thinking disposition and research competence among clinical nurses: A cross-sectional study. *Journal of Clinical Nursing, 29*(7–8), 1332–1340. https://doi.org/10.1111/jocn.15201

Chen, S., & Fan, J. (2022). Validation of the psychometric properties of the Self-Directed Learning Readiness Scale. *Nursing Open, 10*(3), 1639–1646. https://doi.org/10.1002/nop2.1418

Coelho, G. L. H., Vilar, R., Hanel, P. H. P., Monteiro, R. P., Ribeiro, M. G. C., & Gouveia, V. V. (2018). Optimism scale: Evidence of psychometric validity in two countries and correlations with personality. *Personality and Individual Differences, 134*, 245–251. https://doi.org/10.1016/j.paid.2018.06.030

Cui, R., & Zhao, L. (2024). Assessing students' critical thinking in dialogue. *Journal of Intelligence, 12*(11), 106. https://doi.org/10.3390/jintelligence12110106

Dávila-Villavicencio, R., Arias, S., Bedón, A., Perez-Brenis, J., Echabaudes-Ilizarbe, R., & Cunza-Aranzábal, D. (2024). Design and psychometric properties of the student perception of teacher care scale in university students. *Education Sciences, 14*(6), 605. https://doi.org/10.3390/educsci14060605

Emiliannur, E., Hamidah, I., Zainul, A., & Wulan, A. (2017). Using performance assessment model in physics laboratory to increase students' critical thinking disposition. *Journal of Physics: Conference Series, 895*, Article 012143. https://doi.org/10.1088/1742-6596/895/1/012143

Galindo-Domínguez, H., Bezanilla, M., Campo, L., Nogueira, D., & Ruiz, M. (2023). A teachers' based approach to assessing the perception of critical thinking in education university students based on their age and gender. *Frontiers in Education, 8*. 1127705. https://doi.org/10.3389/feduc.2023.1127705

Haeruddin, H., Prasetyo, Z. K., Supahar, S., Sesa, E., & Lembah, G. (2020). Psychometric and structural evaluation of the Physics Metacognition Inventory instrument. *European Journal of Educational Research, 9*(1), 215–225. https://doi.org/10.12973/eu-jer.9.1.215

Jalali, A., Naghibzadeh, A., Rostami, M., Ahmadi, Y., Khodamorovati, M., Vatandost, S., … & Moradi, K. (2024a). Coronary Artery Disease Empowerment Scale (CADES): Persian translation and psychometric properties. *BMC Cardiovascular Disorders, 24*(1). https://doi.org/10.1186/s12872-024-04369-x

Jalali, A., Chavoshani, F., Rasad, R., Darvishi, N., Fashi, F., Khodamorovati, M., … & Moradi, K. (2024b). Transcultural adaptation and psychometric properties of the Persian version of the Nursing Student Competence Scale (NSCS). *SAGE Open Nursing, 10*. https://doi.org/10.1177/23779608241299275

Leach, S. M., Immekus, J. C., French, B. F., & Hand, B. (2019, April 8). The factorial validity of the Cornell Critical Thinking Tests: A multi-analytic approach [Conference paper]. Advancements and Applications in Test Validity Research and Evaluation, Toronto, Canada. https://doi.org/10.3102/1431545

Li, S., Tang, S., Geng, X., & Liu, Q. (2022). Constructing a critical thinking evaluation framework for college students majoring in the humanities. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.1017885

Lins-Kusterer, L., Vieira, N., & Brites, C. (2024). Confirmatory validation of the transgender health care humanization scale. *International Journal for Equity in Health, 23*, 265. https://doi.org/10.1186/s12939-024-02351-9

Luo, S., & Zou, D. (2024). University learners' readiness for ChatGPT-assisted English learning: Scale development and validation. *European Journal of Education, 60*(1). https://doi.org/10.1111/ejed.12886

Maison, M., Syahrial, S., Syamsurizal, S., & Tanti, T. (2019). Learning environment, students' beliefs, and self-regulation in learning physics: Structural equation modeling. *Journal of Baltic Science Education, 18*(3), 389–403. https://doi.org/10.33225/jbse/19.18.389

Manggaberani, A., & Putro, N. H. P. S. (2024). Development of character assessment instrument on English learning for middle school students. *Research and Development in Education (RaDEn), 4*(1), 374–389. https://doi.org/10.22219/raden.v4i1.31923

Mirza, M., Khurshid, K., Shah, Z., Ullah, I., Binbusayyis, A., & Mahdavi, M. (2022). ILS validity analysis for secondary grade through factor analysis and internal consistency reliability. *Sustainability, 14*(13), 7950. https://doi.org/10.3390/su14137950

Moghaddam, R. G. (2019). Development of an instrument to measure EFL teachers' perceptions of reflective teaching. *Indonesian Journal of Applied Linguistics, 9*(1). https://doi.org/10.17509/ijal.v9i1.12826

Oh, J., Kim, E., & Kim, H. (2025). Influence of critical thinking disposition, clinical reasoning competence, and nursing practice environment on medication safety competence of hospital nurses. *Healthcare, 13*(5), 542. https://doi.org/10.3390/healthcare13050542

Papagiannopoulou, T., Vaiopoulou, J., & Stamovlasis, D. (2025). Teachers' readiness to implement robotics in education: Validation and measurement invariance of TRi-Robotics Scale via confirmatory factor analysis and network psychometrics. *Behavioral Sciences, 15*(9), 1227. https://doi.org/10.3390/bs15091227

Roberts, L. D., Heritage, B., & Gasson, N. (2015). The measurement of psychological literacy: A first approximation. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00105

Saputri, M., Indah, R. N., & Rasyid, F. (2022). Debate, critical thinking disposition, and self-confidence: Do they contribute to speaking proficiency? *Premise: Journal of English Education, 11*(1), 189. https://doi.org/10.24127/pj.v11i1.4533

Sireerat, K., Seki, N., Akiyama, M., Kinoshita, A., & Morio, I. (2022). Critical thinking disposition among Thai dental students. *Journal of Dental Education, 86*(8), 968–975. https://doi.org/10.1002/jdd.12913

Stewart, S., & Dempsey, L. (2005). A longitudinal study of baccalaureate nursing students' critical thinking dispositions. *Journal of Nursing Education, 44*(2), 81–84. https://doi.org/10.3928/01484834-20050201-07

Suprapto, N. (2019). Development and validation of students' perception on learning by questioning scale in physics. *International Journal of Instruction, 12*(2), 243–258. https://doi.org/10.29333/iji.2019.12216a

Tep, P., Maneewan, S., & Chuathong, S. (2021). Psychometric examination of Runco Ideational Behavior Scale: Thai adaptation. *Psicologia: Reflexão e Crítica, 34*(1). https://doi.org/10.1186/s41155-020-00170-9

Trigueros, R., Aguilar-Parra, J. M., Sánchez-Iglesias, A., González-Bernal, J. J., & Mercader, I. (2020). Adaptation and validation of the Multi-Dimensional Perceived Autonomy Support Scale for Physical Education to the Spanish physical exercise context. *International Journal of Environmental Research and Public Health, 17*(11), 3841. https://doi.org/10.3390/ijerph17113841

Umar, Ç. (2023). Investigation of critical thinking disposition and reflective thinking skills towards problem-solving of sports sciences students. *The Online Journal of Recreation and Sport, 12*(4), 687–700. https://doi.org/10.22282/tojras.1330861

Xu, L., Abdullah, T., Liu, Q., & Shahroom, N. (2023). Effects of critical thinking disposition on foreign language proficiency in foreign language learning: Evidence from China. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4350357

Yuan, L., Wang, D., Lew, B., Osman, A., & Jia, C. X. (2018). The Future Disposition Inventory-24: Reliability and validity estimates in a large sample of Chinese university students. *BMC Psychiatry, 18*(1). https://doi.org/10.1186/s12888-018-1875-8