Implementation of The Critical Thinking Skills Measurement Instrument Integrated with the Critical and Creative Thinking in Research-Based Learning Model

Mawardi, Yohana Setiawan*, Suhandi Astuti

Universitas Kristen Satya Wacana, Diponegoro St. No.52-60, Salatiga, Central Java, 50711, Indonesia
*Corresponding author, email: yohana.setiawan@uksw.edu

https://doi.org/10.17977/um065.v7.i2.2027.2

Article history
Submitted: 21 July 2026
Revised: 13 August 2026
Accepted: 16 August 2026
Published: 19 August 2026

Keywords
CCTRbL
Critical thinking skills
CTSMI

Abstract
This study addresses the limited availability of critical thinking assessment instruments that comprehensively measure university students' competencies in higher education. Existing FRISCO-based instruments are constrained by suboptimal internal consistency, a predominant focus on cognitive dimensions, limited assessment of critical thinking dispositions, and potential subjectivity arising from inconsistent interpretation of assessment criteria. To overcome these limitations, the Critical Thinking Skills Measurement Instrument (CTSMI) integrated with the Critical and Creative Thinking Research-based Learning (CCTRbL) model was developed. This study aimed to examine the validity, effectiveness, implementation fidelity, and user acceptability of the integrated CTSMI. A mixed-methods approach combining a pre-experimental design and Collaborative Action Research (CAR) was employed, involving four lecturers and a hundred undergraduate students. Expert validation by specialists in psychology, linguistics, and educational evaluation produced mean scores of 77.0, 77.8, and 72.0, respectively, confirming the instrument's validity. Empirical testing yielded a Cronbach's alpha coefficient of 0.734, while Corrected Item–Total Correlation values demonstrated that all items met the validity criterion ($r \geq 0.30$). Implementation of the CTSMI integrated with the CCTRbL model resulted in a mean critical thinking score of 68, indicating a high level of performance. The findings demonstrate that the CTSMI is a valid and reliable instrument for assessing and promoting university students' critical thinking skills. Furthermore, the instrument exhibited strong user acceptability and was implemented effectively, supporting its applicability as a comprehensive assessment tool for higher education learning environments.

1. Introduction

The persistently low levels of university students' critical and creative thinking skills remain a major challenge for higher education in Indonesia in responding to the demands of twenty-first-century learning (Akmam et al., 2019). Several comparative studies indicate that the critical thinking skills of Indonesian university students remain below those of students in several Southeast Asian countries, particularly Malaysia and Singapore (Putri et al., 2023). This condition is influenced by both internal and external factors. Internally, many university students have not yet developed the analytical, evaluative, and reflective thinking patterns that constitute the foundation of higher-order thinking. However, the development of critical and creative thinking skills requires meaningful learning experiences that foster cognitive transformation through strengthening literacy practices, effective learning time management, the utilization of diverse learning resources, a clear understanding of learning objectives, and enhanced motivation and intellectual curiosity (Southworth, 2022). Meanwhile, from an external perspective, higher education practices in Indonesia continue to be dominated by teacher-centered approaches, limiting students' opportunities to develop twenty-first-century competencies, particularly critical thinking, creative thinking, problem-solving, and research skills (Kryshtanovych et al., 2021; Leasa et al., 2021; Sumarni & Kadarwati, 2020).

Consequently, there is a need for instructional models that provide authentic, research-oriented learning experiences. One such model is Critical and Creative Thinking in Research-based Learning (CCTRbL), which was conceptually designed to integrate the development of critical thinking, creative thinking, problem-solving, and research competencies within a unified learning process (Agustin, 2021; Park et al., 2023; Schmaltz et al., 2017). The successful implementation of the CCTRbL model depends not only on the quality of instructional design but also on the availability of assessment instruments capable of both measuring and fostering the development of students' critical and creative thinking skills. Contemporary assessment theory conceptualizes assessment through three complementary functions: assessment of learning, assessment as learning, and assessment for

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

learning (Mawardi, 2018; Mawardi & Satyawati, 2025). Accordingly, assessment instruments should no longer be regarded solely as tools for evaluating learning outcomes but also as mechanisms for promoting learning, encouraging self-reflection, and facilitating the continuous development of students' higher-order thinking skills (Mawardi et al., 2019).

Further analysis of the assessment instrument revealed that the instrument commonly used to evaluate students' critical thinking skills, namely the FRISCO framework (Focus, Reasons, Inference, Situation, Clarity, and Overview), was not fully aligned with the pedagogical characteristics and instructional phases of the Critical and Creative Thinking in Research-based Learning (CCTRbL) model. Moreover, FRISCO-based instruments primarily emphasize cognitive dimensions of critical thinking and therefore do not adequately capture critical and creative thinking dispositions in a comprehensive manner. They also exhibit limitations in terms of reliability and practical implementation, reducing their effectiveness for assessing and supporting the development of higher-order thinking within the CCTRbL learning environment (Asmara et al., 2023).

Reflective evaluation of instructional practice conducted by the research team revealed that the implementation of research-based learning remained predominantly lecturer-centered. Approximately 80% of students reported that they did not fully understand the instructional syntax of the learning model and frequently experienced confusion throughout the learning process. From the lecturers' perspective, 70% indicated that the implementation of the model had not yet reached an optimal level. Consequently, students' critical thinking skills showed no substantial improvement. Empirical evidence from 80 undergraduate students indicated a mean critical thinking score of 66.8, with scores ranging from 44 to 83. This average score suggests a considerable performance gap, corresponding to a 33.2% deficit relative to the expected level of critical thinking proficiency.

In response to these challenges, this study developed the Critical Thinking Skills Measurement Instrument (CTSMI), an assessment instrument specifically aligned with the pedagogical characteristics of the CCTRbL model. The integration of CTSMI with CCTRbL represents a novel contribution because both were developed based on the CCTRbL framework established through the 2024 DRTPM-funded research and further refined through the 2025 DPPM-funded project (Mawardi et al., 2026). The CCTRbL model has been officially registered for intellectual property protection in Indonesia, including copyright for the learning model (No. EC00202230586), the implementation guidebook (No. 000819112), and the CTSMI instrument (No. 001017862). The CCTRbL model is a research-based instructional framework comprising five sequential phases: (1) preliminary and literature study; (2) problem formulation and hypothesis development; (3) data collection and processing; (4) data analysis and interpretation; and (5) contextualization of research findings. In contrast, the CTSMI was constructed based on the theoretical dimensions of critical thinking, encompassing the ability to focus on relevant problems, provide well-founded reasoning, draw contextually appropriate inferences, communicate ideas clearly, and engage in reflective thinking to support decision-making and problem-solving. Based on this theoretical framework, the CTSMI was developed as a five-level analytic rubric, with scores ranging from 5 (highest level of performance) to 1 (lowest level of performance) (see Table 1).

Table 1. The Critical Thinking Skills Measurement Instrument (CTSMI).

Aspects	Rubric and Scoring				
	5	4	3	2	1
Focus	All concepts are correct, clear, and specific.	Most concepts are correct, clear, but less specific.	Some concepts are correct and clear.	Concepts are unfocused, excessive, or doubtful.	All concepts are incorrect or insufficient.
Reason	All statements are supported with strong reasoning.	Most statements are supported with strong reasoning.	Some statements are supported with strong reasoning.	Statements are not supported with strong reasoning.	Statements are not supported with reasoning.
Inference	Conclusions are highly logical, consistent, and based on well-analyzed evidence.	Conclusions are logical and generally supported by evidence.	A conclusion is present, but weak or somewhat forced.	Conclusions are illogical or inconsistent with arguments.	No conclusion or completely misinterpreted.
Situation	Statements and conclusions are analyzed according to context and relevant to arguments.	Statements and conclusions are contextually appropriate but less relevant to arguments.	Statements and conclusions are less contextually appropriate but relevant to arguments.	Statements and conclusions are less contextually appropriate and less relevant to arguments.	Statements and conclusions are not contextually appropriate and irrelevant to arguments.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Aspects	Rubric and Scoring				
	5	4	3	2	1
Clarity	All descriptions of tasks/papers are very clear, systematic, and free of ambiguity.	Most descriptions of tasks/papers are very clear, systematic, and free of ambiguity.	Some descriptions of tasks/papers are very clear, systematic, and free of ambiguity.	Descriptions of tasks/papers are less clear, less systematic, though free of ambiguity.	Descriptions of tasks/papers are unclear, unsystematic, and ambiguous.
Overview	Provides a comprehensive and logical critical reflection.	Provides a comprehensive and logical reflection, but not critical.	Provides a reflective and comprehensive description, but not critical or logical.	Provides a critical reflection, but not comprehensive or logical.	No comprehensive and logical critical reflection is provided.

In response to the identified shortcomings in research-based learning practices, this study aimed to: (1) validate the CTSMI integrated with the CCTRbL model for fostering university students' critical thinking skills; (2) evaluate the effectiveness of the CTSMI; and (3) determine the feasibility of implementing the CCTRbL model using the CTSMI as the primary assessment instrument.

2. Method

This study employed a Collaborative Action Research (CAR) design integrated with an experimental approach as part of the third phase of a modified Research and Development (R&D) framework. The modified R&D framework comprised three sequential phases: (1) preliminary study, (2) model development, and (3) model validation through implementation (Prasanti et al., 2026). The first two phases have been completed, whereas the third phase, namely model validation through implementation, constitutes the focus of the present study. The CAR procedure adopted the Stringer model, as modified by Andanti, Sutikyanto, and Sukisno (Andanti et al., 2024), consisting of three iterative stages: Look, Think, and Act (Figure 1). The Look stage involved systematic observation through data collection to establish a comprehensive understanding of the existing conditions. During this stage, the research identified the key problems, relevant stakeholders, and the contextual factors associated with the issues under investigation. The Think stage comprised critical reflection through the analysis, interpretation, and evaluation of the collected data to identify the root causes of the problems and formulate evidence-based improvement strategies grounded in both theoretical perspectives and empirical findings. The Act stage involved implementing the planned interventions, monitoring their outcomes, and refining the intervention strategies to inform subsequent improvement cycles. The study was conducted over two action cycles, each incorporating an experimental design to examine and compare the statistical significance of the intervention effects between Cycle 1 and Cycle 2.

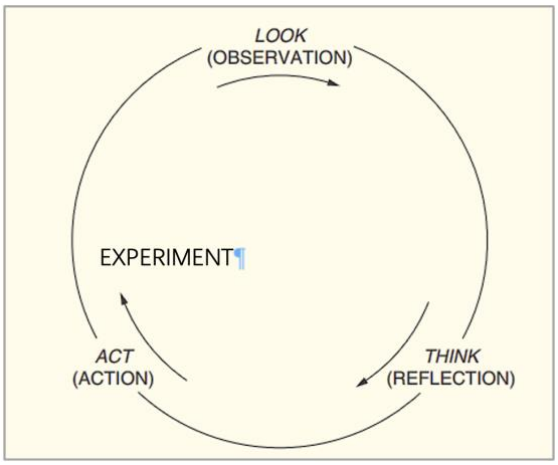


Figure 1. Integration of the Stringer Action Research Model with an Experimental Design.

The participants were undergraduate and graduate students enrolled in the courses Innovative Instructional Design Development for Elementary Civic Education, Fundamental Mathematics for Elementary Education, and Studies and Innovations in Learning during the 2025/2026 academic year. The study involved four lecturers and 100 students from the Primary School Teacher Education undergraduate program, the Bachelor of International Primary Education (BIPE) program, and the Master's Program in Elementary Education.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Data were collected using three primary instruments: 1) an expert validation instrument; 2) the Critical Thinking Skills Measurement Instrument (CTSMI); and 3) an implementation fidelity and acceptability observation sheet, designed to assess lecturers' and students' responses to the implementation of the CTSMI integrated with the Collaborative Critical Thinking Reflective-Based Learning (CCTRbL) model. The scores obtained from expert validation and field implementation were interpreted according to the following predetermined criteria: Very Valid/Critical (≥81); Valid/Critical (61–80); Moderately Valid/Critical (41–60); Less Valid/Critical (21–40); and Very Poor (≤20). To strengthen data credibility through methodological triangulation, structured observation sheets were also employed. To enhance data triangulation, classroom observations were conducted using structured observation sheets. For the Critical Thinking Skills Measurement Instrument (CTSMI), which served as the primary instrument for measuring the main research variable, scores were calculated by dividing the obtained score by the maximum possible score and multiplying the result by 100 to obtain a percentage score.

Data analysis was conducted using several complementary analytical techniques corresponding to the characteristics of each instrument. These included: 1) descriptive categorical analysis to summarize the expert validation results; 2) statistical analyses to examine the validity and reliability of the CTSMI using IBM SPSS Statistics version 26; 3) the Wilcoxon signed-rank test to evaluate differences between paired measurements; 4) quantitative–qualitative descriptive analysis, including percentage distributions and categorical interpretations, to describe students' critical thinking skills as measured by the CTSMI; and 5) qualitative descriptive analysis to examine the implementation fidelity and user acceptability of the CTSMI integrated with the CCTRbL model.

3. Results and Discussion

3.1. Results

This section presents the research findings in accordance with the research problems and objectives, namely: (1) the results of expert validation and the reliability and validity testing of the Critical Thinking Skills Measurement Instrument (CTSMI); (2) the effectiveness of the CTSMI integrated with the CCTRbL model in fostering students' critical thinking skills; and (3) reflections on the implementation of the CTSMI integrated with the CCTRbL model.

3.1.1. Expert Validation and Reliability-Validity Testing of the CTSMI

The expert validation of the CTSMI was conducted by specialists in psychology, language, and educational evaluation. The validation results are summarized in Tables 2, 3, and 4.

Table 2 indicates that the psychology expert validation of the CTSMI yielded a mean score of 77, which falls within the valid category (score range: 61–80). This finding confirms that the rubric-based instrument is appropriate for assessing university students' critical thinking skills. The instrument's validity is supported by six psychological dimensions: psychological attention (70), reasoning (83), inference (77), communication (78), metacognition (77), and decision-making/problem-solving (77). Each dimension reflects the instrument's capacity to measure distinct cognitive and reflective processes that constitute the core components of critical thinking. Among these dimensions, reasoning, encompassing deductive, inductive, and analogical thinking, demonstrated the strongest contribution to the overall adequacy of the instrument. In contrast, psychological attention received the lowest score, suggesting that its indicators require further refinement to more comprehensively assess attentional control. Overall, these findings demonstrate that the CTSMI possesses strong psychometric adequacy for evaluating higher-order cognitive functions underlying critical thinking, particularly reasoning and reflective judgment.

Table 2. Summary of the Psychology Expert Validation Results

No	Psychological Aspect	Validation Score
1.	Psychological Attention Function	70
2.	Reasoning Function	83
3.	Inference Function	77
4.	Communication Function	78
5.	Metacognition Function (Self-Reflection)	77
6.	Decision Making & Problem-Solving Functions	77
Average		77

The linguistic validation results presented in Table 3 show an overall mean score of 77.8, which is categorized as valid. Across the linguistic dimensions, grammar obtained a score of 71.5, vocabulary 80, spelling and punctuation 78.5, discourse coherence 78.5, clarity of ideas 78.5, and appropriateness to language development 80. These findings indicate that the rubric demonstrates adequate linguistic clarity, coherence, and appropriateness for use in higher education contexts, thereby supporting consistent interpretation among

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

raters. Nevertheless, further refinement of the syntactic structure is recommended to improve grammatical accuracy and strengthen inter-rater reliability.

Table 3. Summary of the Language Expert Validation Results

No	Language Aspects	Validation Score
1.	Grammatical (Morphology & Syntax)	71.5
2.	Vocabulary	78.5
3.	Spelling and Punctuation	80
4.	Discourses	78.5
5.	Clarity of Ideas	78.5
6.	Compatibility with Language Development Levels	80
Average		77.8

The results of the educational evaluation expert validation are presented in Table 4. The CTSMI achieved an average score of 72, placing it within the valid category. This indicates that the developed instrument possesses satisfactory quality and is appropriate for practical application. Across the evaluated aspects, construct validity received a score of 70, content validity 80, indicator scalability 70, reliability/consistency 60, freedom from bias 60, usability 80, and relevance to 21st-century competencies 80.

Table 4. Summary of the Educational Evaluation Expert Validation Results

No	Evaluation Aspects	Validation Score
1.	Construct	70
2.	Content Validity	80
3.	Indicator Scalability	70
4.	Reliability / Consistency	60
5.	Bias Free	60
6.	Usability	80
7.	Relevant to 21st Century Competencies	80
Average		72

An empirical pilot study was subsequently conducted to examine the reliability and item validity of the CTSMI among undergraduate students in the Primary School Teacher Education Program and graduate students in the Master of Primary Education Program. Reliability and item validity analyses were performed using SPSS Version 26. A total of 100 students participated in the pilot testing. The pilot study was designed to evaluate the psychometric quality of the rubric-based instrument for measuring critical thinking skills. Statistically, the instrument's adequacy was assessed through Cronbach's alpha reliability coefficient and item validity indices.

The SPSS output indicates a Cronbach's alpha coefficient of 0.734, which falls within the acceptable reliability category. Accordingly, the CTSMI demonstrates satisfactory internal consistency and can be considered a reliable instrument for assessing university students' critical thinking skills.

As shown in Table 5, the internal validity analysis based on the Corrected Item–Total Correlation (CITC) output from SPSS indicates that all instrument items satisfy the established validity criterion, with correlation coefficients of ≥ 0.30. The obtained coefficients ranged from 0.300 to 0.853. Item 3 demonstrated the highest validity coefficient (0.853), indicating a very strong correlation with the total instrument score. This was followed by Item 2 (0.657) and Item 6 (0.502), both of which exhibited good item validity. Meanwhile, Item 1 (0.415), Item 4 (0.340), and Item 5 (0.300) were classified as having acceptable validity, although they remained above the minimum acceptable threshold. Therefore, all six items were deemed valid and retained for measuring the intended research construct.

Table 5. SPSS Output of the Instrument Validity Analysis

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item 1	17.29	3.830	.415	.720
Item 2	16.44	3.787	.657	.647
Item 3	16.71	3.511	.853	.593
Item 4	16.36	4.581	.340	.731
Item 5	16.14	4.907	.300	.734
Item 6	17.27	3.476	.502	.695

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

3.1.2. Effectiveness of the CTSMI Integrated with the CCTRbL Model in Fostering Critical Thinking Skills

The implementation of the CTSMI integrated with the Critical and Creative Thinking-based Learning (CCTRbL) model was conducted in three university courses: Innovative Instructional Design for Primary Civic Education, Fundamental Mathematics for Primary Education, and Studies and Innovations in Learning. The implementation followed Stringer's Action Research Model, as described in the Methods section, with this phase corresponding to the Act stage of the research cycle. The results of the critical thinking assessments conducted during Cycles 1 and 2 using the CTSMI are presented in Tables 6 and 7.

Table 6. Descriptive Statistics of Students' Critical Thinking Skills in Cycles 1 and 2

Statistic	Critical Thinking Skills	
	Cycle 1	Cycle 2
Average	67	68
Min	43	43
Max	80	83

N=100

The descriptive analysis presented in Table 6 indicates an improvement in students' critical thinking skills from Cycle 1 to Cycle 2. In Cycle 1, the mean critical thinking score was 67, which increased to 68 in Cycle 2. Likewise, the maximum score improved from 80 to 83, suggesting a positive development in students' overall critical thinking performance following the implementation of the CTSMI integrated with the CCTRbL model.

Table 7. Frequency Distribution of Students' Critical Thinking Skills in Cycles 1 and 2

Interval	Category	Critical Thinking Skills			
		Cycle 1		Cycle 2	
		f	%	f	%
≥ 80	Very high	2	2	7	7
60-79	High	95	95	88	88
40-59	Medium	3	3	5	5
20-39	Less	0	0	0	0
≤ 19	Very less	0	0	0	0
Total		100	100	100	100

The frequency distribution shown in Table 6 reveals that the majority of students achieved the high category of critical thinking skills, accounting for 95% of participants in Cycle 1 and 88% in Cycle 2. The proportion of students classified in the very high category increased from 2% to 7%, while the percentage of students in the moderate category rose slightly from 3% to 5%. No students were classified in either the low or very low categories during either cycle. Overall, students' critical thinking skills remained predominantly within the high-performance category, accompanied by an increase in the proportion of students achieving the very high level.

To determine whether the observed improvement between Cycles 1 and 2 was statistically significant, a Wilcoxon Signed-Rank Test was performed. The results, presented in Table 8, showed 3 negative ranks, 42 positive ranks, and 55 ties. These findings indicate that only three students experienced a decline in their critical thinking scores, whereas forty-two students demonstrated improvement, and fifty-five students obtained identical scores across both cycles.

Table 8. Wilcoxon Signed-Rank Test Results for Students' Critical Thinking Skills in Cycles 1 and 2

		N	Mean Rank	Sum of Ranks
Cycles_2-Cycles_1	Negative Ranks	3[a]	27.33	82.00
	Positive Ranks	42[b]	22.69	953.00
	Ties	55[c]		
	Total	100		

a. Cycle_2 < Cycle_1
b. Cycle_2 > Cycle_1
c. Cycle_2 = Cycle_1

Test Statistics[a]

	cycle_2-cycle_1
Z	-5.255b
Asymp. Sig. (2-tailed)	.000

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Furthermore, the Test Statistics output reported an Asymp. Sig. (2-tailed) value of 0.000. Since 0.000 < 0.05, a statistically significant difference exists between students' critical thinking scores in Cycle 1 and Cycle 2. These findings demonstrate that the CTSMI integrated with the CCTRbL model effectively promotes the development of university students' critical thinking skills.

Further analysis of the implementation of the CTSMI integrated with the CCTRbL model (Table 9) indicates that the instructional syntax of the CCTRbL model is strongly aligned with all CTSMI indicators. During the first instructional phase (preliminary study and literature review), learning activities primarily fostered the Focus component through the identification and exploration of relevant phenomena. Beginning with the second through fifth phases namely problem and hypothesis formulation, data collection and processing, data analysis and interpretation, and contextualization of research findings all critical thinking indicators (Focus, Reasons, Inference, Situation, Clarify, and Overview) were systematically and simultaneously facilitated. These findings suggest that the integration of the CTSMI with the CCTRbL model supports the comprehensive development of critical thinking skills through structured, contextualized, and research-oriented learning activities.

Table 9. Analysis of the Implementation of the CTSMI Integrated with the CCTRbL Model

Phase of CCTRbL	Impact on critical thinking skills					
	Focus	Reasons	Inferent	Situation	Clarify	Overview
1st phase: Preliminary studies and literature	√	-	-	-	-	-
2nd phase: Problem formulation and hypothesis	√	√	√	√	√	√
3rd phase: Data collection and processing	√	√	√	√	√	√
4th phase: Data analysis and interpretation	√	√	√	√	√	√
5th phase: Contextualization of research findings	√	√	√	√	√	√

3.1.3. Reflection on the Implementation of the CCTRbL Model Using the CTSMI

The reflection on the implementation of the CCTRbL model using the CTSMI provides evidence regarding the feasibility of the instructional model. Model feasibility was examined from the perspective of its perceived usefulness through students' reflective responses collected at each instructional phase. The reflections focused on students' perceptions of the successful and less successful aspects of their learning experiences during each stage of the CCTRbL model. Overall, the reflective findings indicate that the CCTRbL model can be implemented systematically and effectively facilitates research-based learning processes. Nevertheless, several technical and pedagogical challenges were identified throughout its implementation. A visual summary of the students' reflections is presented in Figure 2.

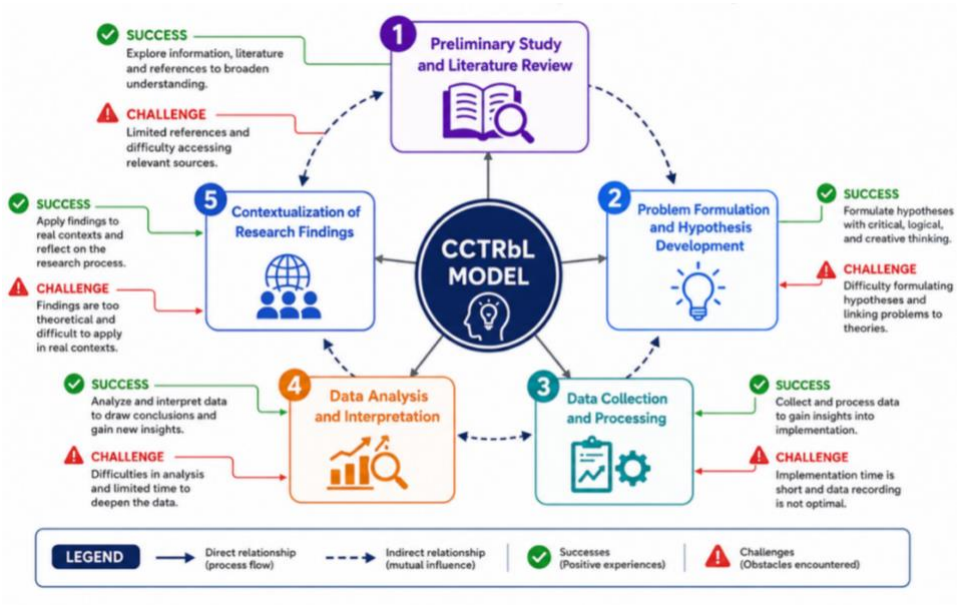


Figure 2. Visualization of Students' Reflections

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

The lecturers' reflections indicate that the implementation of the Critical and Creative Thinking-based Learning (CCTRbL) model progressively fostered the six dimensions of the Critical Thinking Skills Measurement Instrument (CTSMI), namely Focus, Reasons, Inference, Situation, Clarify, and Overview. During Phase 1 (Preliminary Study and Literature Review), instructional activities predominantly developed the Focus dimension. Students demonstrated the ability to identify key issues, define the scope of the problem, and concentrate on relevant information. However, they still required guidance in selecting and evaluating appropriate scholarly references. In Phase 2 (Problem Formulation and Hypothesis Development), all CTSMI dimensions began to be integrated through evidence-based argument construction, hypothesis formulation, contextual analysis, clarification of ideas, and the development of alternative solutions. Nevertheless, some students continued to require support in formulating specific and testable hypotheses. During Phase 3 (Data Collection and Data Processing), all CTSMI indicators were further strengthened through the collection of empirical evidence and systematic data analysis. Despite this progress, students' methodological competence and data analysis skills remained areas requiring further development. Phase 4 (Data Analysis and Interpretation) demonstrated substantial enhancement of the Inference, Reasons, Overview, and Clarify dimensions as students interpreted empirical findings, developed scientifically grounded arguments, and effectively communicated the results of their analyses. Finally, Phase 5 (Contextualization of Research Findings) represented the culmination of the integration of all CTSMI dimensions. At this stage, students were able to relate research findings to authentic instructional practice, engage in comprehensive reflection, and make more adaptive pedagogical decisions based on empirical evidence. These findings suggest that the CCTRbL model provides a systematic framework for cultivating higher-order critical thinking skills through progressive engagement in research-oriented learning activities.

3.2. Discussion

The findings of this study demonstrate that the developed Critical Thinking Skills Measurement Instrument (CTSMI) exhibits satisfactory feasibility from both conceptual and empirical perspectives. Validation by experts in psychology, language, and educational evaluation confirmed that the instrument falls within the valid category. Furthermore, empirical reliability testing yielded a Cronbach's alpha coefficient of 0.734, while all items achieved Corrected Item–Total Correlation (CITC) values above the minimum acceptable threshold of 0.30. These findings indicate that the CTSMI possesses adequate internal consistency and accurately represents the construct of critical thinking skills. They further reinforce the view that the quality of an assessment instrument is determined not only by content validity but also by the appropriateness of its psychological constructs, linguistic clarity, and empirical consistency across items, thereby ensuring stable and trustworthy measurement outcomes. These results are consistent with the theoretical perspectives of Ennis and Facione et al. (Ennis, 1993; Facione et al., 2020) who argue that the validity of a critical thinking assessment instrument depends on its alignment with well-established and scientifically accepted theoretical constructs.

From a linguistic perspective, the validation results indicate that the CTSMI demonstrates satisfactory readability, coherence, and linguistic clarity, thereby enhancing the consistency of interpretation among raters. In rubric-based assessment instruments, clear indicator descriptions are a fundamental prerequisite for achieving high inter-rater reliability. Although the grammar dimension received the lowest score among the linguistic components, it nevertheless remained within the valid category. Consequently, the recommended revisions are primarily editorial rather than substantive in nature. These findings are consistent with those of Gokhale and Almerico (Gokhale & Almerico, 2021) who emphasize that improvements in linguistic quality contribute substantially to enhancing both the reliability and validity of assessment instruments. Nevertheless, revisions to several grammatical features remain necessary to ensure greater syntactic consistency and uniformity throughout the instrument. Overall, the linguistic characteristics of the CTSMI satisfy the requirements for implementation in higher education contexts and are expected to facilitate clearer interpretation and greater measurement accuracy (Shaw et al., 2020). In summary, linguistic quality plays a crucial role in establishing validity in use, particularly when an assessment instrument is employed by multiple evaluators.

The educational evaluation experts further supported the quality of the CTSMI, particularly with respect to its content validity, usability, and relevance to twenty-first-century competencies. In contrast, the reliability/consistency and freedom from bias dimensions received comparatively lower ratings. These findings suggest that instrument development should not conclude with expert validation alone but should be complemented by repeated empirical testing across broader populations to strengthen its psychometric properties. The empirical results of the present study demonstrated that all instrument items satisfied the criteria for internal validity, indicating that the recommendations provided by the expert validators were successfully incorporated during the instrument revision process. Consequently, combining expert validation with empirical testing provides more comprehensive evidence of instrument validity than relying exclusively on either approach. These findings are consistent with the work of Marin and Halpern (Marin & Halpern, 2011) who argue that construct validity depends on the alignment of assessment indicators with their underlying theoretical framework as well as on inter-rater consistency, whereas minimizing bias constitutes a fundamental prerequisite for valid assessment outcomes (Sondergeld & Johnson, 2019). Furthermore, the CTSMI aligns with twenty-first-century learning frameworks by integrating reasoning, metacognition, and problem-solving into

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

the assessment of higher-order thinking skills (Stephenson, 2023). From a theoretical perspective, the instrument conceptualizes critical thinking as a multidimensional construct that integrates analytical and metacognitive dimensions. From a practical perspective, the CTSMI provides a reliable assessment instrument with adequate linguistic clarity that effectively supports the evaluation of university students' higher-order thinking skills (Butler, 2024). To further strengthen its psychometric evidence, future studies should undertake large-scale empirical validation through field testing and Confirmatory Factor Analysis (CFA), accompanied by refinement of the psychological attention indicators and examination of convergent validity through comparisons with established standardized critical thinking instruments (Butler, 2024; Payan-Carreira et al., 2022).

Another significant finding of this study concerns the effectiveness of the rubric-based CTSMI integrated with the Critical and Creative Thinking-based Learning (CCTRbL) model in consistently fostering students' critical thinking skills. Although the increase in the mean score from Cycle 1 to Cycle 2 was relatively modest, the Wilcoxon Signed-Rank Test revealed a statistically significant improvement. This result is consistent with previous studies (Anders et al., 2019; Payan-Carreira et al., 2022) which have demonstrated that rubric-based assessment instruments produce more reliable evaluations when developed from multidimensional and contextually grounded theoretical constructs. In addition, the proportion of students achieving the very high category increased from 2% to 7%, while the majority of students consistently remained within the high category. These findings suggest that the observed improvement reflects not merely an increase in average performance but also a positive shift in the overall distribution of students toward higher competency levels. The statistical significance of this improvement indicates that learning effectiveness is influenced not only by the quality of the assessment instrument but also by the instructional model through which it is implemented. The CCTRbL model actively engages students in problem identification, hypothesis formulation, empirical data collection, evidence analysis, and the contextualization of research findings. Through these sequential activities, students construct knowledge through authentic research experiences. This finding supports constructivist learning theory, which emphasizes that students develop knowledge through interaction with authentic problems and continuous reflective practice tan (Gultom, 2024).

The CTSMI not only demonstrates satisfactory validity and reliability but also provides an integrated assessment of the cognitive, dispositional, and thinking-process dimensions of critical thinking. Consequently, the instrument supports the implementation of both assessment for learning and assessment as learning, while generating comprehensive information regarding students' intellectual development. More importantly, the CTSMI is intended to function not merely as an evaluation tool but as an integral component of the learning process that actively facilitates the continuous development of critical thinking skills. A distinctive contribution of the instrument lies in its integration of the dispositional dimension through reflective evaluation. This component is founded on the premise that critical thinking extends beyond cognitive competence to encompass intellectual dispositions that influence the quality of reasoning and decision-making (Paul & Elder, 2019). Through structured reflective activities, students are encouraged to cultivate intellectual virtues such as open-mindedness, intellectual curiosity, and continuous self-regulation (Brookfield et al., 2019). Consistent with this perspective, numerous studies advocate a holistic assessment approach, in which critical thinking is conceptualized as an externally oriented reasoning process, whereas reflection serves as an internal mechanism for self-evaluation. Such an approach enables a more comprehensive assessment of both the quality of reasoning and the development of students' metacognitive abilities (Binkley et al., 2017; OECD, n.d.). Accordingly, the integration of the CTSMI within the Critical and Creative Thinking-based Learning (CCTRbL) model represents one of the principal novelties of this study, as it accommodates the contextual, reflective, and dispositional dimensions of assessment within a unified framework. Nevertheless, further empirical investigations are warranted to strengthen the instrument's psychometric properties, particularly its construct validity and reliability across broader educational contexts. Overall, this study contributes a theoretically grounded assessment framework that addresses the demands of twenty-first-century learning while providing a robust pedagogical foundation for assessing university students' critical and creative thinking skills.

The analysis of each instructional phase of the CCTRbL model further demonstrates a systematic relationship between learning activities and the six CTSMI dimensions: Focus, Reasons, Inference, Situation, Clarify, and Overview. This finding indicates satisfactory internal coherence, suggesting that these six dimensions collectively provide a coherent representation of the critical thinking construct (Gibson et al., 2025; Yavuz Güler, 2021). The first instructional phase primarily cultivated the Focus dimension through the identification of phenomena and exploration of authentic problems. This phase established the foundation for critical thinking development by encouraging students to recognize relevant issues before engaging in more sophisticated reasoning processes. During the subsequent phases, all CTSMI dimensions developed simultaneously through increasingly complex research activities, including hypothesis formulation, evidence collection, data interpretation, and evidence-based decision-making. This developmental pattern indicates that critical thinking evolves through progressive cognitive development rather than emerging as an instantaneous competency.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

The reflective findings obtained from both lecturers and students further corroborate the quantitative results. Reflections indicated that each instructional phase of the CCTRbL model contributed to the development of different CTSMI dimensions according to the nature of the learning activities. The initial phases primarily emphasized problem identification and focused inquiry, whereas the later phases integrated all dimensions of critical thinking through reflection, interpretation, evaluation, and evidence-based decision-making. Nevertheless, the reflections also identified several challenges, including students' difficulties in selecting high-quality scholarly references, formulating precise hypotheses, and conducting more sophisticated data analyses. These findings suggest that the development of critical thinking requires not only an appropriate instructional model but also systematic instructional scaffolding provided by lecturers, particularly during the early stages of research-based learning implementation (Ennis, 2011; Marin & Halpern, 2011).

Overall, this study offers both theoretical and practical contributions. From a theoretical perspective, it strengthens the argument that the assessment of critical thinking becomes more meaningful when it is directly integrated into the instructional process rather than being employed solely as a summative evaluation tool. From a practical perspective, the CTSMI integrated with the CCTRbL model provides lecturers with an authentic assessment instrument for continuously monitoring the development of students' critical thinking skills throughout the learning process. The integration of the assessment instrument with the instructional syntax enables evaluation to function as assessment for learning, whereby assessment results not only reflect students' learning achievements but also provide evidence to guide subsequent improvements in instructional practice (Mawardi & Satyawati, 2025).

3.3. Implications

The findings of this study suggest that integrating the Critical Thinking Skills Measurement Instrument (CTSMI) with the Critical and Creative Thinking-based Learning (CCTRbL) model represents a promising approach to authentic assessment that not only measures but also facilitates the continuous development of university students' critical thinking skills. These findings underscore the importance of aligning assessment instruments with instructional models so that assessment functions as an integral component of the learning process (assessment for learning) in support of the development of twenty-first-century competencies.

3.4. Limitations

This study has several limitations. First, the implementation and validation of the CTSMI were conducted within a relatively limited sample and instructional context, thereby requiring caution when generalizing the findings to broader educational settings. Second, the psychometric properties of the CTSMI have not yet been examined comprehensively through advanced analyses such as Confirmatory Factor Analysis (CFA), convergent validity testing, or validation across different academic disciplines and institutional contexts. Therefore, further research is recommended to strengthen the evidence supporting the instrument's validity, reliability, and broader applicability.

4. Conclusion

Based on the research objectives and findings presented in this study, the following conclusions can be drawn: The Critical Thinking Skills Measurement Instrument (CTSMI) demonstrated satisfactory validity, as evidenced by expert evaluations from specialists in psychology, language, and educational evaluation, which yielded mean scores of 77, 77.8, and 72, respectively. Furthermore, empirical testing indicated a Cronbach's alpha reliability coefficient of 0.734, while all instrument items achieved empirical validity coefficients (Corrected Item–Total Correlation) of ≥ 0.30, confirming the instrument's acceptable reliability and internal validity. The CTSMI integrated with the Critical and Creative Thinking-based Learning (CCTRbL) model was found to be effective in assessing and fostering students' critical thinking skills, with an average score of 68, corresponding to the high category. The reflective findings indicate that the CCTRbL model, when implemented using the CTSMI, is both feasible and appropriate for supporting research-based learning. The model effectively facilitates the development of students' critical thinking skills through systematic instructional activities, although several technical and pedagogical challenges remain and warrant further refinement in future implementations.

Author Contributions

Mawardi: Conceptualization, Methodology, Formal analysis. Yohana Setiawan: Writing, Data curation, Writing - Original draft preparation. Suhandi Astuti: Project administration, Resources, Investigation.

Funding

This research was funded by Satya Wacana Christian University under the 2026 Fiscal Year Research Grant, based on Master Contract No. 001/SPK-PT/RIK/05/2026, dated May 7, 2026.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration on AI Use

The authors employed artificial intelligence (AI) solely to support language refinement, sentence organization, and the overall clarity of the manuscript. Following the use of these tools, the manuscript underwent thorough review, revision, and verification by the authors. All ideas, interpretations, analyses, and academic contributions presented in this study were developed independently by the authors, who assume full responsibility for the content.

Acknowledgement

The authors would like to express their sincere appreciation and gratitude to Satya Wacana Christian University for its valuable support, which made the successful completion of this research possible.

References

Agustin, A. (2021). Critical and creative thinking for young learners: Asking questions and giving answers in online learning context. *UICELL Conference Proceeding, 5*, 33–40. https://journal.uhamka.ac.id/index.php/uicell/article/view/8287

Akmam, A., Anshari, R., Jalinus, N., & Amran, A. (2019). Factors influencing the critical and creative thinking skills of college students in computational physics courses. *Journal of Physics: Conference Series, 1317*(1), 012172. https://doi.org/10.1088/1742-6596/1317/1/012172

Andanti, M. (2024). Penerapan pembelajaran Agama Buddha berbasis kearifan dan budaya lokal di Sekolah Minggu Buddha (SMB) Adhicitta Suruh, Tengaran, dan Susukan. *Scholaria: Jurnal Pendidikan dan Kebudayaan, 14*(1), 15–25. https://doi.org/10.24246/j.js.2024.v14.i01.p15-25

Anders, P. L., Stellrecht, E. M., Davis, E. L., & McCall, W. D. (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

Asmara, R., Zubaidah, S., Mahanal, S., & Sari, N. (2023). Levels of inquiry and Reading-Questioning-Answering (LoIRQA) to enhance high school students' critical and creative thinking. *International Journal of Instruction, 16*(3), 325–342. https://doi.org/10.29333/iji.2023.16318a

Binkley, D., Campoe, O. C., Alvares, C., Carneiro, R. L., Cegatta, Í., & Stape, J. L. (2017). The interactions of climate, spacing and genetics on clonal Eucalyptus plantations across Brazil and Uruguay. *Forest Ecology and Management, 405*, 271–283. https://doi.org/10.1016/j.foreco.2017.09.050

Brookfield, S. D., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching: An interview with Professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching, 2*(2), 76–90. https://doi.org/10.37074/jalt.2019.2.2.11

Butler, H. A. (2024). Predicting everyday critical thinking: A review of critical thinking assessments. *Journal of Intelligence, 12*(2), 16. https://doi.org/10.3390/jintelligence12020016

Ennis, R. (2011). Critical thinking: Reflection and perspective part II. *Inquiry: Critical Thinking Across the Disciplines, 26*(2), 5–19. https://doi.org/10.5840/inquiryctnews201126215

Ennis, R. H. (1993). Critical thinking assessment. *Theory Into Practice, 32*(3), 179–186. https://doi.org/10.1080/00405849309543594

Facione, P. A., Facione, N. C., & Gittens, C. A. (2020). What the data tell us about human reasoning. In D. Fasko, Jr. & F. Fair (Eds.), *Critical thinking and reasoning: Theory, development, instruction, and assessment* (pp. 272–297). Leiden, the Netherlands: Brill Sense. https://doi.org/10.1163/9789004444591_016

Gibson, C. M., Howard, M. L., & Haight, R. C. (2025). Validation of an instrument to assess student pharmacist self-authorship. *Currents in Pharmacy Teaching and Learning, 17*(2), 102264. https://doi.org/10.1016/j.cptl.2024.102264

Gokhale, A. A., & Almerico, G. M. (2021). Assessing critical thinking in higher education: A systematic review. *Journal of Education and Learning (EduLearn), 10*(5), 22–33.

Gultom, D. A. (2024). Implementasi model pembelajaran berbasis konstruktivisme untuk meningkatkan kemampuan komunikasi matematik dan kemandirian siswa. *Nucleus, 5*(1), 23–28. https://doi.org/10.37010/nuc.v5i1.1509

Kryshtanovych, M., Kryshtanovych, S., Stepanenko, L., Brodiuk, Y., & Fast, A. (2021). Methodological approach to determining the main factors for the development of creative thinking in students of creative professions. *Creativity Studies, 14*(2), 391–404. https://doi.org/10.3846/cs.2021.14806

Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity, 4*(1), 70–76. https://doi.org/10.1016/j.tsc.2009.02.001

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Leasa, M., Batlolona, J. R., & Talakua, M. (2021). Elementary students' creative thinking skills in science in the Maluku Islands, Indonesia. *Creativity Studies, 14*(1), 74–89. https://doi.org/10.3846/cs.2021.11244

Marin, L. M., & Halpern, D. F. (2011). Pedagogy for developing critical thinking in adolescents: Explicit instruction produces greatest gains. *Thinking Skills and Creativity, 6*(1), 1–13. https://doi.org/10.1016/j.tsc.2010.08.002

Mawardi, Agustin, A., & Murti, H. A. S. (2026). An innovative method of measuring students' critical thinking skills in higher education. *Journal of Educational and Social Research, 16*(2), 431–446. https://doi.org/10.36941/jesr-2026-0304

Mawardi, M. (2018). Designing the implementation of model and instructional media. *Scholaria: Jurnal Pendidikan dan Kebudayaan, 8*(1), 26–40. https://doi.org/10.24246/j.js.2018.v8.i1.p26-40

Mawardi, M., Wardani, N. S., Hardini, A. T. A., & Kristin, F. (2019). Model desain pembelajaran tematik terpadu kontekstual untuk meningkatkan kebermaknaan belajar siswa SD. *Scholaria: Jurnal Pendidikan dan Kebudayaan, 9*(1), 48–61. https://doi.org/10.24246/j.js.2019.v9.i1.p48-61

Mawardi, Satyawati, S. T., & Agustin, A. (2025). Fostering twenty-first-century skills: An analysis of the critical and creative thinking in research-based learning model. *The International Journal of Learning in Higher Education, 32*(2), 47–65. https://doi.org/10.18848/2327-7955/CGP/v32i02/47-65

OECD. (2019). An OECD learning framework 2030. In G. Bast, E. G. Carayannis, & D. F. J. Campbell (Eds.), *The future of education and labor* (pp. 23–35). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-030-26068-2_3

Park, J. H., Li, Y., & Niu, W. (2023). Revisiting creativity and critical thinking through content analysis. *Journal of Creativity, 33*(2), 100056. https://doi.org/10.1016/j.yjoc.2023.100056

Paul, R., & Elder, L. (2019). *The nature and functions of critical & creative thinking* (3rd ed.). Lanham, MD: Rowman & Littlefield.

Payan-Carreira, R., Sacau-Fontenla, A., Rebelo, H., Sebastião, L., & Pnevmatikos, D. (2022). Development and validation of a critical thinking assessment-scale short form. *Education Sciences, 12*(12), 938. https://doi.org/10.3390/educsci12120938

Prasanti, A. I., Mawardi, & Winanto, A. (2026). Developing Petualangan Siaga: A TGT-based board game to enhance students' cooperation skills in civic education. *Jurnal Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan, 6*(10), 5. https://doi.org/10.17977/um065.v6.i10.2026.5

Putri, A. S., Prasetyo, Z. K., Purwastuti, L. A., Prodjosantoso, A. K., & Putranta, H. (2023). Effectiveness of STEAM-based blended learning on students' critical and creative thinking skills. *International Journal of Evaluation and Research in Education, 12*(1), 44–52. https://doi.org/10.11591/ijere.v12i1.22506

Schmaltz, R. M., Jansen, E., & Wenckowski, N. (2017). Redefining critical thinking: Teaching students to think like scientists. *Frontiers in Psychology, 8*, 459. https://doi.org/10.3389/fpsyg.2017.00459

Shaw, A., Liu, O. L., Gu, L., Kardonova, E., Chirikov, I., Li, G., … Loyalka, P. (2020). Thinking critically about critical thinking: Validating the Russian HEIghten® critical thinking assessment. *Studies in Higher Education, 45*(9), 1933–1948. https://doi.org/10.1080/03075079.2019.1672640

Sondergeld, T. A., & Johnson, C. C. (2019). Development and validation of a 21st century skills assessment: Using an iterative multimethod approach. *School Science and Mathematics, 119*(6), 312–326. https://doi.org/10.1111/ssm.12355

Southworth, J. (2022). Bridging critical thinking and transformative learning: The role of perspective-taking. *Theory and Research in Education, 20*(1), 44–63. https://doi.org/10.1177/14778785221090853

Stephenson, L. (2023). Collective creativity and wellbeing dispositions: Children's perceptions of learning through drama. *Thinking Skills and Creativity, 47*, 101188. https://doi.org/10.1016/j.tsc.2022.101188

Sumarni, W., & Kadarwati, S. (2020). Ethno-STEM project-based learning: Its impact to critical and creative thinking skills. *Jurnal Pendidikan IPA Indonesia, 9*(1), 11–21. https://doi.org/10.15294/jpii.v9i1.21754

Yavuz Güler, Ç. (2021). The Turkish adaptation of the Triangular Relationship Inventory: The validity and reliability study. *The Family Journal*. Advance online publication. https://doi.org/10.1177/10664807211023002

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174