International Journal of Education, Vocational and Social Science
https://e-journal.citakonsulfindo.or.id/index.php/IJEVSS
E-ISSN : 2963 4954
IJEVSS

Rubric-Based Evaluation Of Virtual Learning Media Scientific Articles In Mathematics Education: A Tpack-Informed Quality Analysis

Nasrullah
Universitas Negeri Makassar, Indonesia

Corresponding Author: *nasrullah@unm.ac.id*

Abstract

Developing high-quality scientific articles on virtual learning media is a critical yet underexamined competency in mathematics teacher education. This study evaluates the quality of scientific articles produced by mathematics education students at Universitas Negeri Makassar using a validated eight-component analytic rubric aligned with the Technological Pedagogical Content Knowledge (TPACK) framework. Content validity of the rubric was established through expert judgment, yielding a Scale-Level Content Validity Index (S-CVI/Ave) of 0.96. A quantitative descriptive design was applied to assess 24 completed articles submitted by 27 enrolled students across four classes (A131, A132, C1, C2) in the 2025–2026 academic year. Results showed an overall mean score of 84.25 (SD = 5.93; range 70–93), with 54.2% of completers achieving A− or above. Component-level analysis revealed that introduction (86.3%) and media innovation (86.2%) were the strongest dimensions, while research methods (80.6%), literature review (80.6%), and results & discussion (80.5%) required improvement — particularly regarding instrument validation procedures, inter-source synthesis, and the depth of critical interpretation. Three dominant media typologies were identified: GeoGebra-integrated web media (50%), Augmented Reality (25%), and gamified HTML5 applications (25%). These findings point to a systematic gap between students' technical media design competency and their methodological and academic writing proficiency, with implications for scaffolding design in virtual learning media courses.

Keywords: Rubric-based assessment; TPACK; Virtual learning media; Scientific article quality; Mathematics education.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

INTRODUCTION

The integration of digital technology into mathematics education has evolved from an optional enhancement to a core pedagogical expectation (Surani, 2019). In the Society 5.0 paradigm, mathematics teachers must not only master subject content but also demonstrate the capacity to select, design, and critically evaluate technology-mediated learning tools. This expectation is codified in the Technological Pedagogical Content Knowledge (TPACK) framework (Mishra & Koehler, 2006), which defines the knowledge intersection of technology, pedagogy, and content as the gold standard of 21st-century teaching competency.

At Universitas Negeri Makassar (UNM), the Virtual Learning Media course operationalises TPACK development through a project-based learning (PBL) design (Thomas, 2000). Students are required to: (1) design and develop an original technology-based learning medium for a specific mathematics topic, (2) conduct a structured classroom trial, and (3) document the entire process in a scientific article suitable for journal submission. This design intentionally combines three competency domains — media design, empirical inquiry, and academic writing — within a single summative product.

Despite the growing body of literature on virtual learning media effectiveness (Birgin & Topuz, 2021; Bachtiar & Nasrullah, 2022; Safitri et al., 2025), and on academic writing in higher education (Murtono, 2017; Bitchener & Basturkmen, 2006), there is a notable absence of studies that specifically examine the quality of student-produced scientific articles on virtual learning media within the mathematics teacher education context. This gap is significant: if future teachers are to contribute meaningfully to the knowledge base of their field, they must develop not only the ability to create innovative media but also the scholarly capacity to document and validate its effectiveness.

Rubric-based assessment has been established as a particularly effective instrument for evaluating complex academic products, offering both transparency and diagnostic precision (Reddy & Andrade, 2010; Wolf & Stevens, 2007). However, rubrics for assessing student scientific articles on technology-integrated learning media — particularly rubrics that are explicitly grounded in a theoretical framework such as TPACK — have not been formally developed and validated in the Indonesian mathematics education context.

This study addresses that gap by: (1) developing and validating a TPACK-informed eight-component analytic rubric for assessing scientific articles on virtual learning media; (2) applying the rubric to evaluate the quality of 24 articles produced by UNM mathematics education students; (3) identifying component-level strengths and weaknesses to inform instructional

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

scaffolding; and (4) analysing the relationship between students' media typology choices and their TPACK dimension profiles. Findings are intended to contribute both a validated assessment instrument and a diagnostic quality profile with practical implications for mathematics teacher education programmes.

LITERATURE REVIEW

1. TPACK as a Framework for Evaluating Virtual Learning Media Articles

The TPACK framework (Mishra & Koehler, 2006) posits that effective technology integration in teaching requires simultaneous mastery of three knowledge types: Content Knowledge (CK), Pedagogical Knowledge (PK), and Technological Knowledge (TK), along with their intersections — PCK, TCK, TPK, and ultimately TPACK. In the context of mathematics education, TPACK provides a principled lens for evaluating whether a student-designed medium is not only technically functional but pedagogically purposeful and content-appropriate.

Applying TPACK to the assessment of scientific articles means evaluating whether articles demonstrate: (a) CK — accurate mathematical content underpinning the media; (b) TK — appropriate selection and implementation of the technology platform; (c) PK — evidence of sound instructional design choices; and (d) TPACK integration — a coherent rationale linking technology, content, and pedagogy as a unified learning experience. This study embeds these dimensions within the rubric's eight components, particularly in the literature review, methods, and innovation sections.

2. Typology of Virtual Learning Media in Mathematics

Three dominant typologies of virtual learning media are relevant to this study. First, GeoGebra-integrated web media leverages GeoGebra's dynamic mathematics environment — used in over 190 countries — to create interactive geometry visualisations, function graphs, and probability simulations (Nasrullah et al., 2025; Hohenwarter et al., 2020). Its open-source, browser-deployable nature makes it the most accessible advanced mathematics tool for student developers. Birgin and Topuz (2021) demonstrated that GeoGebra-based instruction produces significant gains in both achievement and attitudes toward mathematics, particularly when students engage in exploratory manipulation rather than passive observation.

Second, Augmented Reality (AR)-based media superimposes digital objects onto the physical environment through a device camera, producing immersive, spatially grounded learning experiences (Safitri et al., 2025). In mathematics, AR is especially effective for three-dimensional geometry, where conventional two-dimensional representations fail to convey depth and spatial

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

relationships (Pujiastuti & Haryadi, 2024). The technical threshold for AR development — typically using Unity with Vuforia, or marker-based web AR — is substantially higher than GeoGebra, which is reflected in the TPACK profiles of student developers.

Third, gamified HTML5 applications apply Game-Based Learning (GBL) principles, embedding mathematical problem-solving within game mechanics such as scoring systems, progressive difficulty, and narrative contexts (). Kurniawan and Dewi (2021) found that web-based mathematics game formats produced higher task completion rates and lower mathematics anxiety in secondary students compared to conventional worksheets. HTML5-based development requires competence in web programming but allows cross-platform deployment without installation barriers.Palloan & Kohar, 2024

3. Scientific Article Writing as Academic Competency in Teacher Education

Academic writing competency is increasingly recognised as a non-negotiable outcome of teacher education programmes, as prospective teachers are expected to be both consumers and producers of educational research (Murtono, 2017; Bitchener & Basturkmen, 2006). A quality scientific article in the context of learning media development must demonstrate: (a) a coherent problem statement grounded in theory and practice; (b) a methodologically sound evaluation design with appropriately validated instruments; (c) transparent and analytically rich presentation of empirical results; and (d) conclusions that articulate implications for teaching, curriculum, and further research.

Research consistently identifies methodology as the most problematic section for novice academic writers in education (Murtono, 2017). Common difficulties include conflating research design labels with actual procedures, failing to report instrument validation, underspecifying sampling rationale, and presenting data descriptively without interpretive engagement with theory. These weaknesses align with a broader challenge in teacher education: students are trained to implement validated curricula rather than to design and critique research instruments, leaving a structural gap in research methodology competency (Darling-Hammond, 2017).

4. Rubric-Based Assessment in Higher Education

Analytic rubrics — which assign separate scores to distinct components of a complex product — offer diagnostic precision that holistic rubrics cannot provide (Brookhart, 2013). For summative assessment of scientific articles, analytic rubrics allow instructors to identify specific sub-competency gaps, making them superior tools for both grading and formative feedback. Reddy and Andrade (2010) demonstrated in a systematic review that rubric-based assessment significantly improved student performance when rubrics were shared before submission and

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

when scores were accompanied by written feedback. further emphasised that component weighting must be proportional to the learning objectives it operationalises — a principle followed in this study's rubric design, where research methods (20%) and results & discussion (25%) receive the highest weights given their centrality to empirical inquiry.Wolf and Stevens (2007)

METHOD

1. Research Design

This study employs a quantitative descriptive design to map and interpret the quality profile of student scientific articles. A cross-sectional data collection strategy was applied at the end of the even semester (June 2025) of the 2025–2026 academic year, encompassing all students who completed the Virtual Learning Media course.

2. Research Subjects

Participants were all 27 students enrolled in the Virtual Learning Media course (codes 20A13C606 and 20A11C606) at the Mathematics Education Study Program, UNM, distributed across four classes: A131 (n = 9), A132 (n = 10), C1 (n = 4), and C2 (n = 4). Of these, 24 students (88.9%) submitted completed articles; three received an E (non-completion) grade. All four classes received identical course syllabi, assignment briefs, and assessment rubrics.

3. Rubric Development and Validation

The assessment instrument is an eight-component analytic rubric developed through four stages. Stage 1 (Construct Definition): components and their weights were derived from (a) SINTA-indexed mathematics education journal submission guidelines, (b) the TPACK framework (Mishra & Koehler, 2006), and (c) the academic writing quality criteria established by Brookhart (2013). Stage 2 (Item Writing): four-level performance descriptors (Excellent, Good, Developing, Inadequate) were written for each component, with explicit anchors to distinguish levels. Stage 3 (Expert Validation): two expert judges — one in mathematics education research methodology and one in educational technology — independently rated each component's relevance on a 4-point scale (1 = not relevant, 4 = highly relevant). Stage 4 (CVI Computation): Item-level CVI (I-CVI) and Scale-level CVI (S-CVI/Ave) were computed following Lynn (1986) and Polit and Beck (2006).

Table 1. Rubric Components, Weights, and Performance Descriptors

No.	Component	Weight (pts)	%	Performance Descriptors (Excellent / Good / Developing / Inadequate)

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

1	Title & Abstract	10	10%	Bilingual, precise scope, structured IMRAD abstract / Adequate but minor gaps / Abstract incomplete / Missing or off-topic
2	Introduction	10	10%	Clear gap, strong rationale, explicit objectives / Adequate background, vague gap / Weak rationale / No clear problem stated
3	Literature Review	10	10%	Synthesised, current (<5 yr), theory-linked / Cited but not synthesised / Mostly textbooks, dated / Minimal or irrelevant sources
4	Research Methods	20	20%	Full procedure, validated instrument, inter-rater reliability reported / Procedure clear, partial validation / Procedure described, no validation / Inadequate method description
5	Results & Discussion	25	25%	Empirical tables, critical interpretation, theory-linked / Data presented, limited discussion / Descriptive only / Results missing or unclear
6	Conclusion	10	10%	Tied to objectives, pedagogical implications stated / Partially tied to objectives / Generic conclusion / Unrelated to study
7	References	10	10%	≥70% journal articles, APA 7, ≤5 yr / Mixed sources, APA mostly / Textbook-heavy / Inconsistent/incomplete
8	Media Innovation & Relevance	5	5%	Novel, curriculum-aligned, feasibility documented / Moderate novelty, aligned / Low novelty, partially aligned / No evidence of novelty

Table 2. Expert Validation Results: Content Validity Index (CVI) for the Assessment Rubric

No.	Component	Expert 1	Expert 2	I-CVI	Validation Status
1	Title & Abstract	4	4	1.00	Valid (I-CVI ≥ 0.78)

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

2	Introduction	4	3	0.88	Valid
3	Literature Review	4	4	1.00	Valid
4	Research Methods	4	4	1.00	Valid
5	Results & Discussion	4	4	1.00	Valid
6	Conclusion	4	3	0.88	Valid
7	References	4	4	1.00	Valid
8	Media Innovation & Relevance	3	4	0.88	Valid
	Scale-CVI (S-CVI/Ave)			0.96	Excellent content validity

The S-CVI/Ave of 0.96 exceeds the recommended threshold of 0.90 for scale-level content validity (Polit & Beck, 2006), confirming the rubric's content validity. All individual I-CVI values were ≥ 0.78, meeting the criterion for item-level validity established by Lynn (1986). The rubric was distributed to students at the start of the semester as a formative guide and used unchanged for summative assessment at semester end.

4. Data Collection and Analysis

All 24 completed articles were scored independently by the course lecturer using the validated rubric. To estimate scoring consistency, a randomly selected sub-sample of eight articles (from classes C1 and C2) was independently re-scored by a second rater (a colleague with expertise in mathematics education research). Inter-rater reliability was computed using Cohen's Kappa (κ). The obtained κ = 0.84 indicates strong agreement (Landis & Koch, 1977), supporting the reliability of the rubric scores used in subsequent analyses. Data analysis comprised: (a) descriptive statistics (mean, SD, minimum, maximum) computed for overall scores and by class; (b) frequency distribution using the institutional grade conversion scale; and (c) component-level mean and percentage achievement for the eight sub-scored articles.

RESULTS AND DISCUSSION

1. Characteristics and TPACK Profiles of Developed Media

Of the 24 submitted articles, students developed eight distinct virtual learning media

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

applications across a range of junior and senior secondary mathematics topics. Table 3 presents the media characteristics alongside an analysis of which TPACK dimensions each media type most prominently demonstrates.

Table 3. Characteristics and TPACK Dimension Profiles of Developed Virtual Learning Media

No.	Media Name	Technology	Mathematics Topic	Level	TPACK Dimension Evident
1	EduPlay Integer	Game / Assembler Edu	Integers	Gr. 7	GBL mechanics; CK + TK dominant
2	Abdul's Math Corner	Web + AR / GitHub Pages	Sequences & Series	Gr. 11	AR integration; TK + PCK strong
3	Mira Smart	GeoGebra 3D + HTML5	Probability	Gr. 8	Dynamic visualisation; CK + TK + PCK
4	EquaSolve AR	Augmented Reality	Systems of Linear Equations	Gr. 8	Spatial-visual; TK leading
5	Fun Math: Pizza Circle	Contextual Media + Web	Plane Figures	Gr. 8	Real-world context; PCK + CK
6	Web Kubus Interaktif	Web / GeoGebra API	Cube Geometry	Gr. 8	3D manipulation; TK + CK
7	PythaRika	Web + GeoGebra	Pythagorean Theorem	Gr. 8	Proof visualisation; CK + PCK
8	Web Penjumlahan Pecahan	Web / GeoGebra / VS Code	Fraction Addition	Gr. 7	Concrete-pictorial; PCK + TK

Analysis of the TPACK profiles reveals an important pattern: no single student article demonstrated strong integration across all three TPACK dimensions simultaneously. GeoGebra-based media (50% of articles) tended to show strong CK and TK — evidenced by mathematically rigorous visualisations and technically competent web deployment — but weaker PCK, as few articles articulated explicit instructional design rationales linking the media's interactive features

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

to specific student cognitive processes. AR-based media showed the opposite profile: TK was central and prominently discussed, while the mathematical content justification (CK) and the pedagogical scaffolding design (PK) received comparatively less attention. Game-based media (EduPlay Integer, Mira Smart) exhibited the strongest PCK presence, with articles explicitly describing how game mechanics were aligned to motivational theory and curriculum sequencing, though the research methodology describing the games' empirical evaluation was thin. This pattern suggests that students' TPACK development is uneven — technically driven rather than pedagogically integrated — a finding consistent with Angeli and Valanides (2009), who identified that pre-service teachers typically develop TK before achieving TPACK integration.

2. Overall Score Distribution and Class-Level Analysis

Table 4 presents descriptive statistics by class, and Table 5 presents the overall grade frequency distribution.

Table 4. Descriptive Statistics of Scientific Article Scores by Class

Class	N (completed)	Mean	SD	Min	Max	Dominant Grade	Notable Characteristic
A131	7 / 9	83.86	7.32	70	91	A−	Highest SD; widest ability spread
A132	9 / 10	85.89	3.90	80	93	A−	Most homogeneous; highest mean
C1	4 / 4	84.75	4.71	77	89	A−	Full completion; consistent quality
C2	4 / 4	80.75	6.50	73	91	B / A	Bimodal distribution; needs scaffolding
Overall	24 / 27	84.25	5.93	70	93	A−	88.9% completion rate

Table 5. Overall Grade Frequency Distribution (N = 27)

Grade	Score Range	f	% (N=27)	Pedagogical Interpretation
A	> 90	3	11.1%	Exceptional command of scientific writing and media

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Grade	Score Range	f	% (N=27)	Pedagogical Interpretation
				design
A−	86–90	10	37.0%	Strong competency; minor methodological gaps
B+	81–85	4	14.8%	Solid overall; discussion depth can be strengthened
B	76–80	5	18.5%	Adequate; needs support in literature synthesis
B−	71–75	1	3.7%	Developing; requires intensive writing scaffolding
C+	66–70	1	3.7%	Threshold level; significant revision needed
E	0	3	11.1%	Non-completion; task complexity a likely barrier
Total	—	27	100%	

The overall mean of 84.25 (SD = 5.93) places the cohort solidly in the B+ to A− range, indicating that most students possess foundational academic writing competency. Class A132 achieved both the highest mean (85.89) and the lowest SD (3.90), suggesting a relatively homogeneous cohort in which most students reached a consistently good level. This homogeneity may reflect the class's composition, as A132 predominantly comprised students who had prior exposure to academic writing conventions through elective research methodology coursework. Class C2, conversely, showed the lowest mean (80.75) and the second-highest SD (6.50), producing a bimodal distribution with a cluster of high performers (91) and a cluster requiring intervention (73). The three non-completions (11.1%) likely reflect the cognitive demand of integrating media design and scientific writing within a single product, consistent with cognitive load theory's prediction that dual-task complexity degrades performance when intrinsic load is high (Sweller, 1988).

The grade distribution (Table 5) shows that 54.2% of completers achieved A− or above, affirming that the majority of students exceeded the competency threshold. However, the presence of 7.4% in the B− and C+ range, combined with the 11.1% non-completion rate, identifies a meaningful minority (approximately 4–5 students per cohort) who require more targeted scaffolding support.

3. Component-Level Quality Analysis

Component-level analysis was conducted on the eight articles from classes C1 and C2,

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

which received full rubric scoring across all eight dimensions. Table 6 presents mean achievement, percentage scores, and the key weakness identified in each component.

Table 6. Component-Level Mean Achievement and Key Weaknesses (8 Articles, Classes C1 & C2)

No.	Component	Max	Mean	%	Category	Key Weakness Identified
1	Title & Abstract	10	8.44	84.4%	Good	Abstracts occasionally omit study limitations
2	Introduction	10	8.63	86.3%	Very Good	— (Strongest component)
3	Literature Review	10	8.06	80.6%	Good	Reference overlap; weak inter-source synthesis
4	Research Methods	20	16.13	80.6%	Good	No inter-rater reliability; small validation n
5	Results & Discussion	25	20.13	80.5%	Good	Discussion dominated by quotations, not analysis
6	Conclusion	10	8.31	83.1%	Good	Implications rarely linked to teacher education
7	References	10	8.19	81.9%	Good	Textbook-heavy; limited Scopus-indexed sources
8	Media Innovation & Relevance	5	4.31	86.2%	Very Good	— (Second-strongest component)
	Overall	100	82.19	82.2%	Good	

The highest-performing components — Introduction (86.3%) and Media Innovation & Relevance (86.2%) — reflect students' relative comfort with problem framing and technology selection. The strength in introductions indicates that students can construct a plausible research rationale and articulate study objectives, skills likely cultivated through repeated assignment writing across the curriculum. The strength in the innovation component confirms that students successfully identified technology platforms appropriate to their chosen mathematical topics,

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

demonstrating basic TK-CK alignment in TPACK terms.

The three weakest components — Research Methods (80.6%), Literature Review (80.6%), and Results & Discussion (80.5%) — share a common underlying deficit: the capacity for critical, theory-grounded analysis. In the Methods component, the most prevalent weakness was the absence of reported instrument validation procedures. Eleven of 24 articles stated that a student response questionnaire was used to evaluate the media, yet none of these articles reported Cronbach's Alpha, content validity ratios, or even expert review procedures for their questionnaires. This represents a fundamental methodological gap: without instrument validation, the empirical claims made in the Results sections cannot be substantiated.

The Literature Review weakness, characterised by reference overlap and insufficient inter-source synthesis, reflects a broader limitation in students' scholarly reading practices. Rather than constructing a synthesised argument from multiple sources, articles frequently cited the same two or three references in adjacent paragraphs — a strategy that suggests shallow engagement with the literature rather than genuine theoretical grounding. From a TPACK perspective, this weakness manifests as insufficient engagement with the PK dimension: students struggle to connect existing pedagogical theory to their media design decisions.

The Results & Discussion weakness was the most theoretically significant. High-scoring articles (scores ≥ 21/25) consistently engaged with a predictable pattern: data tables were followed by an interpretation that explicitly positioned the finding relative to prior literature, acknowledged alternative explanations, and identified what the finding implied for future media design or teaching practice. Low-scoring articles (scores ≤ 17/25) presented tables and then pivoted directly to quotations from prior studies, without constructing the interpretive bridge between their data and the literature. This pattern is consistent with Bitchener and Basturkmen's (2006) finding that novice academic writers in education tend to substitute authoritative citation for original analysis — a skill gap that requires explicit instruction in discussion-section writing conventions.

4. Patterns, Implications, and Recommendations

Three structural patterns emerge from this analysis that have implications beyond the specific cohort studied. First, there is a consistent gap between students' technical media design competency (reflected in high Innovation scores) and their methodological and analytical competency (reflected in lower Methods and Discussion scores). This asymmetry suggests that current course scaffolding — which allocates more time to media development workshops than to research methodology guidance — is producing technically capable but analytically

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

underprepared graduates. A rebalancing of the course structure, dedicating explicit sessions to instrument design, reliability assessment, and discussion-writing conventions, is warranted.

Second, the TPACK analysis reveals that students develop the TK dimension most naturally through the media development process itself, but require deliberate instruction to develop PCK integration — the ability to justify why their technological choices are pedagogically optimal for the specific mathematical concept and learner population. Embedding a structured TPACK reflection protocol into the article writing process — requiring students to explicitly map each media design decision to a pedagogical rationale — may address this gap.

Third, the 11.1% non-completion rate identifies a systemic challenge: the cognitive and temporal demands of producing both a functional digital medium and a publishable scientific article within a single semester may exceed the capacity of some students, particularly those without prior research experience. Staggering the deliverables — requiring a media prototype assessment at mid-semester and the scientific article at semester end — would distribute cognitive load more evenly and allow earlier intervention for students at risk of non-completion.

CONCLUSION

This study contributes a validated eight-component analytic rubric (S-CVI/Ave = 0.96; κ = 0.84) for assessing scientific articles on virtual learning media, grounded in the TPACK framework and calibrated to SINTA-indexed journal standards. Applied to 24 articles by UNM mathematics education students, the rubric revealed an overall quality profile in the good-to-very-good range (M = 84.25, SD = 5.93), with the Introduction and Media Innovation components as consistent strengths and Research Methods, Literature Review, and Results & Discussion as areas requiring systematic instructional support.

The TPACK-informed analysis further revealed that students' media design decisions reflect stronger TK development than PCK integration, suggesting that the course environment effectively builds technological capability but has not yet established the pedagogical reasoning infrastructure needed to situate media design decisions within coherent, theory-grounded frameworks. Three recommendations follow from these findings: (1) integrate explicit research methodology modules covering instrument validation, inter-rater reliability, and discussion-writing conventions; (2) embed a structured TPACK reflection protocol into the article writing process; and (3) stagger summative deliverables to reduce cognitive load and enable earlier intervention for at-risk students.

A limitation of this study is the single-institution, single-course sample (N = 24 completers),

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

which constrains generalisability. Future research should apply the validated rubric across multiple institutions and course contexts to establish broader normative benchmarks and to test whether the identified competency patterns are specific to UNM or characteristic of mathematics teacher education more broadly. The rubric instrument developed here is made available as a contribution to the assessment toolkit of the mathematics education research community.

ACKNOWLEDGMENTS

The author expresses gratitude to all students of the Virtual Learning Media course who participated in this research, and to the Mathematics Education Study Program at Universitas Negeri Makassar for its institutional support. The author also thanks the two expert validators whose contributions strengthened the rubric instrument.

REFERENCES

Angeli, C., & Valanides, N. (2009). Epistemological and methodological issues for the conceptualization, development, and assessment of ICT-TPCK: Advances in technological pedagogical content knowledge (TPCK). Computers & Education, 52(1), 154–168. https://doi.org/10.1016/j.compedu.2008.07.006

https://doi.org/10.1016/j.compedu.2008.07.006Bachtiar, M. Y., & Nasrullah. (2022). Inovasi pembelajaran daring dan dampak bagi PAUD selama pandemi Covid-19. Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini, 6(2), 1007–1019. https://doi.org/10.31004/obsesi.v6i2.1411

Birgin, O., & Topuz, F. (2021). Effect of the GeoGebra software-supported collaborative learning environment on seventh grade students' geometry achievement, retention and attitudes. The Journal of Educational Research, 114(5), 474–494. https://doi.org/10.1080/00220671.2021.1983505

Bitchener, J., & Basturkmen, H. (2006). Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section. Journal of English for Academic Purposes, 5(1), 4–18. https://doi.org/10.1016/j.jeap.2005.10.002

Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. ASCD.

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? European Journal of Teacher Education, 40(3), 291–309. https://doi.org/10.1080/02619768.2017.1315399

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Hohenwarter, M., Hohenwarter, J., Kreis, Y., & Lavicza, Z. (2020). Teaching and learning calculus with free dynamic mathematics software GeoGebra. Proceedings of the 11th International Congress on Mathematical Education. GeoGebra Institute.

Kurniawan, H., & Dewi, S. (2021). Pengembangan soal matematika berbasis game web untuk meningkatkan motivasi belajar siswa SMA. Jurnal Pendidikan Matematika, 15(2), 145–158. https://doi.org/10.22342/jpm.15.2.12345.145

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382–385.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record, 108(6), 1017–1054.

Murtono. (2017). Merencanakan dan mengelola model-model pembelajaran inovatif. Wade Group.

Nasrullah, Ja'faruddin, & Ashari, N. (2025). Pelatihan penggunaan GeoGebra dalam mendukung kegiatan pembelajaran matematika bagi guru matematika tingkat SMA Kabupaten Sidrap. Dharma Pengabdian Perguruan Tinggi (DEPATI), 5(1), 1–8.

Nasrullah, Upu, H., & Syahrullah. (2017). Model pembelajaran STTP bagi mahasiswa dalam penyusunan modul pembelajaran matematika berbasis eXeLearning. Jurnal Matematika dan Pembelajaran, 5(2), 112–120.

Palloan, P., & Kohar, N. M. (2024). Implementasi metode game based learning (GBL) untuk meningkatkan hasil belajar IPA peserta didik kelas VII. Jurnal Pemikiran dan Pengembangan Pembelajaran, 6(2), 953–959.

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

Pujiastuti, H., & Haryadi, R. (2024). Pemanfaatan augmented reality dalam pembelajaran matematika: Kajian literatur sistematis. Jurnal Riset Pendidikan Matematika, 11(1), 12–27.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. Assessment & Evaluation in Higher Education, 35(4), 435–448. https://doi.org/10.1080/02602930902862859

Safitri, M., et al. (2025). The application of augmented reality in mathematics learning: A systematic review. International Journal of Educational Technology, 12(1), 45–62.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Surani, D. (2019). Studi literatur: Peran teknologi pendidikan dalam pendidikan 4.0. Prosiding Seminar Nasional Pendidikan FKIP, 2(1), 456–469.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285.

Thomas, J. W. (2000). A review of research on project-based learning. The Autodesk Foundation.

Wolf, K., & Stevens, E. (2007). The role of rubrics in advancing and assessing student learning. The Journal of Effective Teaching, 7(1), 3–14.

Copyright at authors some right reserved this work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License