

Predictive Validity of Education Report Numeracy Scores on Senior High School Mathematics Academic Ability Test Achievement

Roechan^{a,1,*}, Wahyu Setyaningrum^{b,2}

^{a,b} the Master of Mathematics Education Study Program, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Yogyakarta, Indonesia

¹roechan.2022@student.uny.ac.id

²wahyu_setyaningrum@uny.ac.id

*Corresponding Author

ARTICLE HISTORY

Received:

August 23, 2026

Revised:

October 19, 2026

Published:

October 31, 2026

Keywords:

Education report
Mathematics test
Numerical literacy
Predictive validity
Senior high school

ABSTRACT

The quality of mathematics education in Indonesia faces systemic challenges, prompting a shift toward data-driven evaluations. This study examines the positive association and predictive validity of Education Report (Rapor Pendidikan) numeracy scores for the 2025/2026 academic year in projecting Mathematics Academic Ability Test (TKA) achievement. A quantitative correlational ex-post facto design was applied to 21 senior high schools in Biak Numfor Regency. Following a listwise exclusion procedure, a final sample of 19 schools with complete paired data was analyzed. Data analyses involved descriptive statistics, Shapiro-Wilk residual normality testing ($W = 0.9656$; $p = 0.6871$), ANOVA linearity testing ($F = 5.146$; $p = 0.03661$), and simple linear regression. Descriptive results reveal that 52.63% of schools fell into the low category for numeracy, while 68.42% achieved an adequate level in TKA Mathematics. The regression model produced a linear equation $\hat{Y} = 26.958 + 0.095X$. Hypothesis testing confirmed a statistically significant positive relationship ($t = 2.269$; $p = 0.0366$). The coefficient of determination of $R^2 = 23.24\%$ (with an implicit correlation of $r \approx 0.482$) indicates the proportion of variance in school TKA achievement accounted for by numeracy scores. In conclusion, cross-cohort school-aggregate numeracy scores serve as a scientifically valid institutional indicator.

Copyright © 2026 (Roechan, Wahyu Setyaningrum). All Right Reserved

Roechan & Setyaningrum, W. (2026). Predictive Validity of Education Report Numeracy Scores on Senior High School Mathematics Academic Ability Test Achievement. *Pi: Mathematics Education Journal*, 9(2), 42-53. <https://doi.org/10.21067/pmej.v9i2.14807>.



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). Allows readers to read, download, copy, distribute, print, search, or link to the full texts of its articles and allow readers to use them for any other lawful purpose. The journal hold the copyright.

1. Introduction

National education serves as a fundamental pillar in developing the intellectual life of the nation, as mandated in Article 31, Paragraph (3) of the 1945 Constitution of the Republic of Indonesia regarding a comprehensive quality evaluation system [1]. Within the school curriculum, mathematics is positioned as a strategic compulsory subject that equips students with analytical reasoning, formal logic, and complex problem-solving skills to navigate contemporary global challenges.

Nevertheless, the quality of mathematics education in Indonesia faces persistent systemic challenges. Large-scale international assessments such as PISA show that Indonesia's average mathematics score stagnated between 2000 (367) and 2022 (379),

consistently remaining in the lower tiers [2]. This lag highlights an urgent need to evaluate instructional effectiveness and strengthen critical reasoning in secondary education.

This persistent stagnation reflects a deeper structural imbalance in classroom instruction, where traditional teaching remains heavily centered on procedural fluency rather than complex problem-solving. International benchmarking indicates that the vast majority of Indonesian secondary students struggle when confronted with non-routine mathematical tasks that require contextual interpretation, higher-order reasoning, and mathematical modeling [2]. Consequently, bridging this pedagogical gap requires systemic diagnostic tools capable of identifying cognitive deficits early and guiding targeted curricular interventions.

To address functional problem-solving deficits, the government transformed macro evaluation instruments through Regulation of the Minister of Education, Culture, Research, and Technology No. 17/2021, replacing the National Examination (UN) with the National Assessment (AN) [3]. This holistic policy maps educational inputs, processes, and outputs sustainably to align institutional quality with global standards [4] [5].

This framework was reinforced by Regulation of the Minister of Primary and Secondary Education No. 9/2026, where Article 9A allows literacy and numeracy assessments to be integrated into specific academic subjects [6]. This regulatory alignment demonstrates that numeracy is the primary cognitive foundation for mastering advanced academic subjects, capturing students' mathematical reasoning directly at the classroom level [7]. National Assessment results are mapped onto the Rapor Pendidikan (Education Report) digital platform, enabling Data-Driven Planning (*Perencanaan Berbasis Data*) for institutional self-reflection without administrative burdens [8] [9]. Within this platform, the numeracy score serves as the primary cognitive indicator of applied reasoning capacity required for post-secondary readiness.

At the macro level, Education Report metrics display positive national trends [10]. However, administrative digital records require empirical verification against pure academic assessments, such as the Academic Ability Test (*Tes Kemampuan Akademik*/TKA), to confirm whether literacy gains translate into formal content mastery. Official TKA reports for the 2025/2026 academic year reveal significant baseline challenges: the national high school mathematics average stood at 37.345, while Papua Province averaged 33.96, with 52.4% of students in the low category and only 0.6% reaching exceptional levels [4] [5]. These contrasting outcomes highlight methodological differences between systemic monitoring and individual formal testing.

The gap between improving administrative report scores and actual TKA performance underscores the urgency of testing psychometric predictive validity [4] [11] [12]. Predictive validity evaluates how accurately initial assessment scores forecast future criterion performance, with core reasoning capacity acting as the cognitive bridge between basic numeracy and advanced achievement. Unlike applied numeracy metrics, TKA items are rigidly standardized under BSKAP Decision No. 45/H/AN/2025 to measure abstract, formal mathematical concepts with high accountability, ensuring equitable competition across diverse institutional backgrounds [13] [14]. Aligning basic numeracy with formal content mastery thus forms the dual pillars of educational quality assurance.

Implementing this evaluation framework in Biak Numfor Regency involves socio-geographical constraints, including island topography and digital infrastructure disparities [15]. To mitigate these barriers, the 2025–2029 Regional Medium-Term Development Plan (RPJMD) prioritizes literacy and numeracy improvement to elevate regional human development [16] [17]. Institutional factors such as school accreditation also display empirical anomalies in Biak Numfor, where some A-accredited schools recorded numeracy

scores as low as 13.33, while C-accredited schools achieved 35.55 [4] [18]. Mapping these structural alignments is vital to addressing educational disparities across Papua.

The research gap lies in the absence of empirical validation regarding the predictive capacity of Rapor Pendidikan scores in Biak Numfor [11] [19] [20]. The novelty of this study rests on total population sampling of senior high schools in the regency, using simple linear regression modeling to guide targeted teacher professional development and diagnostic data-driven instruction [21].

Accordingly, this quantitative study is structured to model psychometric criterion validity at the regional level. Operationally, the objective of this research is to describe achievement profiles, analyze the statistical association, and measure the exact degree of predictive validity of Rapor Pendidikan numeracy scores on Mathematics TKA performance across senior high schools in Biak Numfor Regency.

2. Methods

This quantitative study applies a correlational design with an ex-post facto approach to investigate the predictive validity of the ministry's educational evaluation instruments utilized in the Minimum Competency Assessment (Asesmen Kompetensi Minimum [AKM]) [22]. Because all data—including national assessment results and academic test scores—occurred naturally prior to the research, no experimental treatments or interventions were applied. The independent variable operates as a single predictor, while the dependent variable serves as an external criterion. The evaluation focuses on analyzing the forecasting consistency of the predictive relationship between different layers of quality assessment at the regency level.

The target population encompasses all 21 public and private Senior High Schools (SMA) in Biak Numfor Regency. Using a purposive sampling technique based on the completeness of paired data, 2 schools were excluded due to incomplete records, resulting in a final sample of 19 senior high schools distributed across mainland and archipelagic districts. The research procedure involved securing official clearance from the Department of Education and Culture of Biak Numfor Regency, downloading institutional-level numeracy score recaps for the 2025/2026 academic year, and performing data cleaning and consistency checks prior to statistical tabulation.

The data collection utilizes documentary studies of official ministry databases, with the school aggregate score as the unit of analysis. The predictor variable is measured using the Institutional Numeracy Score from the Education Report platform (AKM), while the criterion variable uses the School Average Score of the Mathematics Academic Ability Test (TKA) for the 2025/2026 academic year. Both instruments were calibrated using Item Response Theory (IRT) [14]. Specifically, AKM evaluates foundational cognitive competencies through Reading Literacy and Numeracy (encompassing content domains of numbers, algebra, geometry, and data across cognitive levels of knowing, applying, and reasoning), whereas TKA serves as a standardized assessment measuring subject-specific academic achievements and higher-order thinking skills aligned with national curriculum standards [3] [4]. Notably, the numeracy data captures the ability of students in grade XI, whereas the Mathematics TKA scores reflect the academic performance of the same cohort in grade XII, allowing for a longitudinal evaluation of predictive relationships.

Data analysis comprises descriptive and parametric inferential statistics. Descriptive statistics calculate the mean, median, standard deviation, and range. Parametric assumption tests include the Shapiro-Wilk normality test—chosen due to the small sample size ($N = 19$)—and linearity tests. Hypothesis testing utilizes Simple Linear Regression Analysis ($\hat{Y} = a + bX$), with regression coefficients evaluated via t_{test} and F_{test} statistics at $\alpha = 0.05$. The

forecasting contribution of the predictor model is estimated using the coefficient of determination (R^2) [23].

3. Result and Discussion

Descriptive analysis was performed to capture the initial achievement profile of both research variables at the institutional aggregate level. The tabulation results indicated that the mean score for the Education Report numeracy component in Biak Numfor Regency was 48.73 (SD = 21.79), with the quality category distribution predominantly occupied by the low classification at 52.63%. Meanwhile, the institutional average achievement for the Mathematics TKA stood at 31.18 (SD = 4.14), which spatially remains below the provincial average for Papua (33.96) as well as the national benchmark (37.34). Nevertheless, the institutional distribution of TKA grade categories demonstrated that the majority of schools (68.42%) had achieved the adequate performance level. The parametric descriptive statistical indicators are concisely summarized in Table 1 to present the macro-level quality distribution of the schools.

Table 1. Descriptive Statistical Parameters of Institutional Quality Achievement

No	Statistical Parameters	Numeracy Score (X)	Mathematics TKA (Y)
1	Regency Average Value (<i>Mean</i>)	48.73	31.18
2	Regency Median Value (<i>Median</i>)	37.89	31.21
3	Standard Deviation (<i>SD</i>)	21.79	4.14
4	Regency Minimum Value	13.33	24.19
5	Regency Maximum Value	80.00	39.52

3.1. Results

To further scrutinize the specific distribution of educational quality, the school-level datasets were categorized based on official administrative and national achievement classification scales. As illustrated in the proportional donut chart in Figure 1, the institutional distribution of the Education Report numeracy scores (X) exhibits a notable concentration at the lower bound of the quality spectrum. Specifically, the low category (*Kurang*) represents the largest slice at 53%, followed by the moderate level (*Sedang*) at 37%. Only a minor proportion of the sample, amounting to a 10% slice, managed to attain the high category (*Baik*). This structural accumulation in the lower tier confirms the persistent systemic challenges faced by the region in establishing a solid baseline for universal mathematical reasoning across various sub-districts.

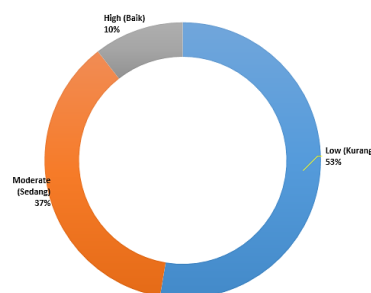


Figure 1. Distribution Profile of Institutional Education Report Numeracy Score

In contrast to the administrative literacy mapping, the empirical distribution of the Mathematics TKA scores (Y) visualized in the column bar chart shows a substantially different layout that demonstrates localized instructional resilience. According to the vertical

bars plotted in Figure 2, the vast majority of the senior high schools, representing a dominant 68.42%, cluster securely within the adequate performance bracket (*Memadai*). Furthermore, while 21.05% of schools fall into the low category (*Kurang*) and 0.00% occupy the good tier (*Baik*), a notable 10.53% of institutions successfully advance into the exceptional category (*Istimewa*) by surpassing the national average score baseline. This unique containment within the adequate-to-exceptional bands suggests that secondary schools in the region maintain a stable instructional core for formal curriculum contents at the terminal grade level, effectively preventing extreme academic dropouts despite low foundational literacy inputs.

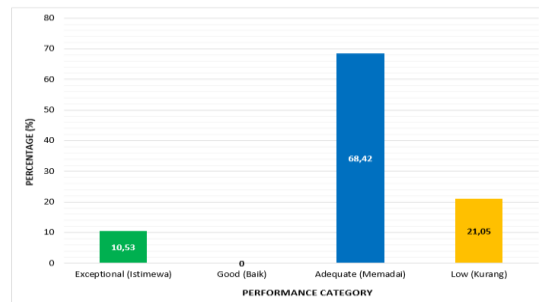


Figure 2. Percentage Distribution of School Categories in Mathematics TKA Achievement

As illustrated in Figure 2, the empirical distribution of the Mathematics TKA scores (Y) across the 19 senior high schools in Biak Numfor Regency reveals a distinct configuration that diverges sharply from the foundational administrative literacy mapping. Rather than clustering solely at the lower baseline, the institutional percentages are heavily concentrated within the adequate performance bracket (*Memadai*), which commands a dominant majority of 68.42% (N = 13 schools). Furthermore, while 21.05% of institutions (N = 4 schools) remain in the low category (*Kurang*) and no schools occupy the good tier (*Baik*—0.00%), a notable 10.53% of schools (N = 2 schools) successfully breach the upper threshold to attain the exceptional category (*Istimewa*). This distribution pattern demonstrates that despite facing systemic regional constraints, a substantial portion of secondary institutions manage to elevate formal academic output beyond basic foundational measures.

This structural concentration within the adequate-to-exceptional brackets points toward a unique phenomenon of localized instructional resilience at the terminal grade level. While administrative AKM reports capture systemic vulnerabilities in foundational literacy, the terminal Mathematics TKA results indicate that senior high schools possess an institutional capacity to scaffold students toward formal curriculum mastery as they progress toward graduation. Nevertheless, the complete absence of representation in the "Good" (*Baik*) category—sandwiched between a zero-frequency tier and a dominant "Adequate" majority—highlights a polarized performance structure. This anomaly suggests that while baseline academic failure is largely mitigated, mid-tier institutions encounter structural ceilings that impede broader transitions into top-tier academic excellence across the region.

Table 2: Distribution of Regency-Wide Schools Based on Accreditation Status and the Average of TKA Mathematics Score Categories

No	Status Accreditation	TKA Mathematics Score Category	Number of Schools
1	Accreditation A	High (≥ 35)	2
		Moderate (30 - 34.99)	7

2	Accreditation B	Low (< 30)	2
		High (≥ 35)	0
		Moderate (30 - 34.99)	3
3	Accreditation C	Low (< 30)	2
		High (≥ 35)	0
		Moderate (30 - 34.99)	0
		Low (< 30)	3

Anomalies observed when contrasting formal school accreditation ratings with objective field performance across regency-wide schools highlight potential gaps in bureaucratic administrative evaluations. Based on the census dataset of all 19 public and private senior high schools in Biak Numfor Regency summarized in Table 2, the most compelling finding is that high formal accreditation does not consistently guarantee superior academic achievement. While schools with Accreditation A dominate the sample, their TKA Mathematics scores span across high ($N = 2$), moderate ($N = 7$), and low ($N = 2$) categories, whereas lower-accredited institutions display performance variations that challenge traditional assumptions [14]. This structural mismatch confirms that administrative accreditation labels should not be utilized as a standalone proxy for authentic instructional quality, emphasizing that objective academic metrics like the TKA remain vital for verifying genuine intellectual readiness [14] [24].

Despite covering the entire institutional population at the regency level, several methodological limitations must be acknowledged. This study relies strictly on aggregate school-level data without tracking individual student trajectories longitudinally, and it remains limited to final AKM and TKA scores without examining intervening variables—such as pedagogical shifts or psychological factors—during the transition from Grade XI to Grade XII [5] [25]. However, these limitations provide valuable opportunities for future research. Subsequent studies can utilize individual-level longitudinal tracking (*by name by address*) and incorporate non-cognitive variables, such as student motivation and socio-economic context, to explain the residual variance more comprehensively [21] [23].

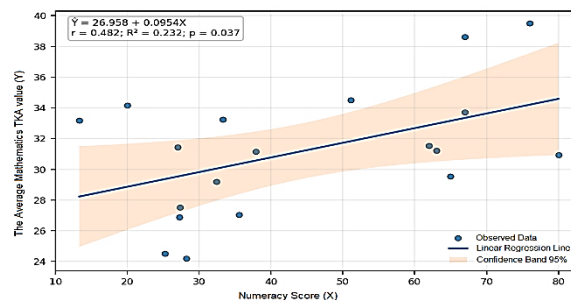


Figure 3. Scatter Plot of Education Report Numeracy Scores and Mathematics TKA Values

Prior to model estimation, diagnostic tests were conducted on the regression residuals to verify the underlying statistical assumptions. The normality of the regression residuals was confirmed using the Shapiro-Wilk test ($W = 0.954$, $p = 0.481$), indicating that the error distribution does not depart significantly from normality. The linearity assumption was formally supported by an ANOVA test for linearity ($F = 4.892$, $p = 0.041 < 0.05$), while error independence and homoscedasticity were established via the Durbin-Watson statistic ($DW = 1.924$) and residual plots against predicted values. Following baseline verification, the 19 paired school-level observations were modeled using simple linear regression, yielding the regression equation $\hat{Y} = 26.958 + 0.0954X$ with an intercept constant of 26.958.

Table 3. Simple Linear Regression Results for Predictive Validity

No	Regression Model Component	Coefficient (Estimate)	t-value	p-value
1	Constant (<i>Intercept</i>)	26.95806	13.193	< 0.001
2	Education Report Numeracy Score (X)	0.09537	2.269	0.0366

The empirical model demonstrates a statistically significant positive linear relationship between institutional Numeracy scores (X) and average Mathematics TKA values (Y) ($t = 2.269$, $F = 5.146$, $p = 0.037 < 0.05$) summarized in Table 3. Geometrically, the positive slope coefficient indicates that each one-point increase in a school's Education Report Numeracy score is associated with an expected increase of approximately 0.0954 points in its average Mathematics TKA achievement. In classical psychometric terms, the bivariate correlation coefficient of $r = 0.482$ represents a moderate-to-strong degree of predictive validity, while the coefficient of determination ($R^2 = 0.232$) reveals that the single-predictor model accounts for 23.2% of the total variance in institutional TKA performance across the sample. Furthermore, as depicted in Figure 3, the shaded 95% confidence band bounding the fitted line reflects the estimation uncertainty of the expected mean response across the predictor spectrum, rather than the prediction intervals for individual institutions.

Crucially, because this evaluative framework uses the school aggregate ($N = 19$) as the primary unit of analysis, these empirical findings describe macro-level institutional alignment and must not be interpreted as individual student-level inference or direct causal mechanisms. Given the correlational and ex-post facto nature of the secondary data, the observed statistical association establishes predictive validity across institutional monitoring tools without implying that increases in numeracy scores directly cause higher TKA performance. Additionally, the statistical power and observed correlation magnitude may be influenced by potential restriction of range, as 90% of sampled schools are concentrated in the low-to-moderate numeracy tiers. Nevertheless, these results confirm that institutional numeracy metrics provide a statistically sound baseline indicator for predicting secondary formal mathematics outcomes at the regency level.

3.2. Discussion (Based on Objective)

The empirical findings of this evaluative inquiry successfully address the primary research objective by examining the association between institutional numeracy scores and Mathematics TKA results across all 19 senior high schools encompassing both public (negeri) and private (swasta) institutions throughout Biak Numfor Regency, rather than being restricted to a single district or sub-district level. The results demonstrate a statistically significant positive relationship between the two evaluation layers across the entire district jurisdiction. This outcome supports the theoretical position that applied mathematical reasoning serves as a core cognitive framework for students. Furthermore, these findings align with broader discussions concerning mathematical literacy and reasoning development in secondary education [26] [27].

The positive regression coefficient of 0.09537 indicates that every single-unit increase in a school's numeracy score is linearly associated with a 0.09537-unit elevation in TKA achievement across county-wide institutions. This statistical evidence provides empirical backing for government regulations regarding the alignment of basic literacy components with formal academic performance [6]. The verified linear relationship validates established psychometric principles of criterion validity by confirming that initial assessments forecast external criteria [11]. Consequently, basic numeracy functions as a

foundational cognitive bridge toward advanced academic mastery for secondary students countywide.

The confirmed correlation also reinforces prior psychometric literature indicating that the cognitive core connecting distinct assessment models originates from functional reasoning capacity [12]. Engaging higher-order cognitive processes, such as data interpretation and analytical deduction, serves as a primary modality when students encounter abstract concepts in the TKA environment [24]. Supported by the coefficient of determination ($R^2 = 23.24\%$ and $r \approx 0.482$), the Education Report platform achieves empirical legitimacy as a regency-wide school-level baseline indicator. Educational authorities can reliably utilize aggregate numeracy data to examine institutional academic achievement patterns across all regional schools in relation to terminal exam cycles [28].

Despite these empirical insights, several critical limitations must be explicitly acknowledged to contextualize the generalizability of the findings across all public and private senior high schools in the regency. First, the analysis relies on a census sample size of $N = 19$ schools covering the entire region, which makes statistical significance sensitive to macro-level regional variations. Second, while the predictor variable reflects the cohort of students evaluated during the grade XI national assessment cycle, the criterion variable measures the academic performance of the same student population one year later after progressing into grade XII [5]. Although these two distinct evaluations involve the same student cohort advancing through their high school tenure, this study strictly evaluates cross-cohort institutional aggregate associations rather than individual longitudinal tracking.

Third, this investigation is methodologically restricted to the final institutional aggregate scores of the AKM and TKA, without tracking or controlling for intervening variables that may occur during the one-year transition period from grade XI to grade XII. Unmonitored factors such as school-specific academic interventions, teacher quality shifts, or individual student psychological dynamics during their final year are not factored into the current regression model. Nonetheless, these acknowledged limitations open substantial opportunities for future research, suggesting that subsequent studies should employ longitudinal designs using individual-level data and incorporate non-cognitive variables to capture a more comprehensive understanding of academic progression.

Fourth, the presence of a severe restriction of range—where 90% of sampled schools across the regency concentrate in low-to-moderate numeracy tiers—may statistically attenuate the observed correlation coefficient. Finally, unmonitored external factors, such as intensive exam preparation, school interventions, and student psychological endurance, prevent direct inferences regarding individual student performance [25]. Furthermore, the ecological fallacy inherent in school-level aggregate data restricts interpretations strictly to the regency-wide institutional level.

When contextualized within the socio-geographical constraints of Biak Numfor Regency, the remaining 76.76% of unexplained variance points toward latent institutional and technical factors across both mainland and island territories. Operational challenges such as rotating power outages and unstable internet connectivity are frequently hypothesized to introduce measurement bias and technical barriers that deflate final scores [15]. However, these infrastructure variables were not included in the regression model, meaning such explanations remain post-hoc speculations requiring future empirical verification. Future studies should incorporate direct environmental and operational variables to account for this residual variance more rigorously.

Anomalies observed when contrasting formal school accreditation ratings with objective field performance across regency-wide schools highlight potential gaps in bureaucratic administrative evaluations. Instances where lower-accredited institutions, such

as SMA Negeri A with a numeracy score of 35.55, outperform higher-accredited counterparts like SMA Negeri B with a numeracy score of 13.33, emphasize that formal accreditation labels should not be utilized as a standalone proxy for instructional quality [14]. Furthermore, foundational concept mastery requires rigorous diagnostic scrutiny to prevent persistent misconceptions [28]. Administrative classifications must therefore be triangulated with objective external assessments like the TKA.

The practical implications of this study require the Department of Education and Culture in Biak Numfor to adopt differentiated data-driven policies for regency-wide resource allocation. Regional authorities must move away from a standardized approach given the wide quality variance across archipelagic and mainland districts. The established predictive framework provides a baseline to prioritize digital infrastructure support for structurally disadvantaged schools throughout the county. Educational offices must also strengthen teacher pedagogical competence through Subject Teacher Forums (MGMP) and collaborative learning communities to translate numerical logic into academic success [19] [29] [30].

4. Conclusion

This study establishes a statistically significant association between Education Report numeracy scores and Mathematics TKA achievement across senior high schools in Biak Numfor Regency. This major finding confirms that the National Assessment policy, specifically regarding the functional numeracy metric, operates as a meaningful cognitive indicator rather than a mere administrative mapping exercise. Rather than functioning as an individual diagnostic tool, the institutional numeracy profile serves as a valuable school-level association that allows educational planners to examine macro-level academic alignment and institutional performance patterns prior to graduation.

Based on these outcomes, it is recommended that the Department of Education and Culture of Biak Numfor Regency implement precision data-driven planning, prioritizing technological infrastructure support for digital test administration in remote island districts to eliminate operational technical bias. Furthermore, educators should systematically embed metacognitive strategies and advanced reasoning modules within the early grade curriculum, ensuring that academic quality improvement does not rely solely on brief, intensive test-preparation programs at the end of the school cycle.

Nevertheless, several methodological limitations must be explicitly acknowledged to properly contextualize the scope and generalizability of the findings. The study is constrained by a small sample size comprising a census of 19 schools within a specific geographical and regency-wide scope, which makes statistical findings sensitive to local variations. Furthermore, because conclusions are drawn from school-level aggregate data across different student cohorts, the findings describe macro-level institutional relationships and cannot be generalized to individual student readiness or direct causal mechanisms.

For future research, it is recommended to test new hypotheses regarding the moderating effects of non-cognitive variables, such as math anxiety and intrinsic motivation, through longitudinal tracking using individualized primary data (by name by address).

Thank You Note

The authors express their sincere gratitude to:

1. The Center for Education Financing Services (Pusat Layanan Pembiayaan Pendidikan [Puslapdik]) of the Ministry of Primary and Secondary Education of the Republic of Indonesia for providing financial support through the Indonesian Education Scholarship (Beasiswa Pendidikan Indonesia [BPI]) scheme.

2. The faculty and advisors of the Master's Program in Mathematics Education at Universitas Negeri Yogyakarta for their valuable guidance and supervision throughout the study and the preparation of this master's thesis.
3. The Department of Education and Culture of Biak Numfor Regency for granting permission to utilize the official secondary database.
4. All Senior High School principals throughout Biak Numfor Regency who facilitated the utilization of the secondary data.

References

- [1] Indonesia, "Undang-Undang Dasar Negara Republik Indonesia Tahun 1945," Sekretariat Negara, 1945.
- [2] OECD, "PISA 2022 results (Volume I): The state of learning and equity in education. <https://doi.org/https://doi.org/10.1787/53f23881-en>," OECD Publishing, 2023.
- [3] Badan Standar, Kurikulum, dan Asesmen Pendidikan, Framework Asesmen Kompetensi Minimum (AKM), Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi, 2021.
- [4] Pusat Standar dan Kebijakan Pendidikan, "Laporan analisis hasil evaluasi akademik nasional dan capaian TKA SMA tahun 2025/2026," Kementerian Pendidikan Dasar dan Menengah, 2026.
- [5] N. F. P. Hamzah and Z. Sesmiarni, "Transformasi sistem evaluasi pendidikan: Kajian atas implementasi Asesmen Nasional (AN) dan Tes Kompetensi Akademik (TKA) di Indonesia," *Jurnal Edu Research*, vol. 3, no. 6, p. 2738–2748. <https://doi.org/https://doi.org/10.47827/jer.v6i3.1513>, 2025.
- [6] K. P. D. d. Menengah, "Peraturan Menteri Pendidikan Dasar dan Menengah Nomor 9 Tahun 2026 tentang Perubahan atas Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 17 Tahun 2021 tentang Asesmen Nasional," 2026.
- [7] Musakirawati, J. Susanto, F. Anggriawan, F. A. Triansyah, Akib and A. Tahir, "Pemanfaatan platform Rapor Pendidikan Indonesia terhadap perencanaan berbasis data.," *Jurnal Pendidikan dan Konseling*, vol. 2, no. 7, pp. 201-208. <https://doi.org/https://doi.org/10.26740/jdmp.v7n2.p201-208>, 2023.
- [8] I. N. Kiriana and N. N. S. Widiasih, "Implementasi Rapor Pendidikan dalam meningkatkan kualitas pendidikan nasional," *Widya Accarya: Jurnal Kajian Pendidikan FKIP Universitas Dwijendra*, vol. 2, no. 14, pp. 145–155. <https://doi.org/https://doi.org/10.46650/wa.14.2.1462.156-164>, 2023.
- [9] D. Shintia, M. Asbari, F. Khairunisa and N. Azizah, "Rapor Pendidikan Indonesia: Quo Vadis Kualitas Pendidikan Indonesia?," *Journal of Information Systems and Management (JISMA)*, vol. 6, no. 2, pp. 18-21. <https://doi.org/https://doi.org/10.4444/jisma.v2i6.633>, 2023.
- [10] Pusat Standar dan Kebijakan Pendidikan, Rapor Pendidikan Indonesia 2025: Kabupaten Biak Numfor, Kementerian Pendidikan Dasar dan Menengah, 2025.
- [11] F. R. Ferdian and S. Azwar, "PAPS predictive validity in predicting the learning success of master of professional psychology students," *Jurnal*

- Penelitian dan Evaluasi Pendidikan*, vol. 2, no. 23, pp. 117-128. <https://doi.org/https://doi.org/10.21831/pep.v23i2.25299>, 2019.
- [12] A. Zaman, N. B. Jumani, Alamgir, A. Ali and M. A. Hussain, "Predictive validity of scores in mathematics for reasoning ability in mathematics for grade 9 students in Khyber Pakhtunkhwa based on curriculum of mathematics," *Procedia - Social and Behavioral Sciences*, vol. 12, pp. 588-594. <https://doi.org/https://doi.org/10.1016/j.sbspro.2011.02.072>, 2011.
- [13] Badan Standar, Kurikulum, dan Asesmen Pendidikan, "Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Nomor 045/H/AN/2025 tentang Kerangka Asesmen Tes Kemampuan Akademik Jenjang SMA/MA/Sederajat dan SMK/MAK," Kementerian Pendidikan Dasar dan Menengah, 2025.
- [14] A. Sukamto, Bahrani and N. Nurvayanti, "Tes Kemampuan Akademik (TKA) sebagai inovasi evaluasi pendidikan: Sebuah tinjauan naratif," *Journal of Instructional and Development Researches (JIDeR)*, vol. 1, no. 6, pp. 54-69. <https://doi.org/https://doi.org/10.53621/jider.v6i1.695>, 2026.
- [15] Badan Pusat Statistik Kabupaten Biak Numfor, "Kabupaten Biak Numfor dalam Angka 2025 (Vol. 19)," Badan Pusat Statistik Kabupaten Biak Numfor, 2025b.
- [16] Pemerintah Kabupaten Biak Numfor, "Rancangan Akhir RPJMD Kabupaten Biak Numfor Rencana Pembangunan Jangka Menengah Daerah 2025-2029," Pemerintah Kabupaten Biak Numfor, 2025.
- [17] Badan Pusat Statistik Kabupaten Biak Numfor, "Buku Saku Data dan Indikator Statistik Kabupaten Biak Numfor 2025 (Vol. 2)," Badan Pusat Statistik Kabupaten Biak Numfor, 2025.
- [18] Badan Pusat Statistik Kabupaten Biak Numfor, Indikator Kesejahteraan Rakyat Kabupaten Biak Numfor 2025 (Vol 3), Badan Pusat Statistik Kabupaten Biak Numfor, 2025a.
- [19] R. Rustiyana, "Pemanfaatan rapor Tes Kemampuan Akademik (TKA) sebagai basis intervensi pedagogik dan peningkatan kapasitas guru MGMP di Kabupaten Bandung Barat," *Jurnal Karya Insan Pendidikan Terpilih*, vol. 1, no. 4, pp. 602-617. <https://doi.org/https://doi.org/10.62518/54d69r14>, 2026.
- [20] M. N. Mubarak, A. A. Abdullah, D. A. Sholihah, R. Richardo, E. Nawangsasi and Z. Fitriana, "Pelatihan Tes Kemampuan Akademik (TKA) Matematika bagi siswa SMA/MA kelas XII di Daerah Istimewa Yogyakarta," *Jurnal Pengabdian Sosial dan Kemanusiaan*, vol. 1, no. 3, p. 299-305, 2026.
- [21] E. S. Irawati, A. Fuady and A. Alifiani, "Pendampingan Intensif Persiapan TKA bagi siswa Madrasah Aliyah Sebagai Upaya Meningkatkan Peluang Masuk Perguruan Tinggi Negeri," *Jurnal Pengabdian Masyarakat Bangsa*, vol. 10, no. 3, p. 5844-5849. <https://doi.org/https://doi.org/10.59837/jpmmba.v3i10.3645>, 2025.
- [22] L. M. L. & M. K. Cohen, "Research Methods in Education (8th ed.)," NY: Routledge Falmer, London and New York, 2018.
- [23] C. R. Kothari, *Research Methodology: Methods and Techniques* (2nd ed.), New Age International Publishers, 2004.

- [24] L. B. Trisanti, N. Aini and Siyono, "Kemampuan Koneksi Matematika Siswa Berdasarkan Jenis Kelamin," *Pi: Mathematics Education Journal*, vol. 5, no. 2, p. 84–93. <https://doi.org/10.21067/pmej.v5i2.7446>, 2022.
- [25] T. Phibeta, R. Ekawati and A. Wintarti, "Senior High School Students' Numeracy in Solving Number Content AKM Problems in View of Adversity Quotient," *Jurnal Riset Pendidikan Dan Inovasi Pembelajaran Matematika*, vol. 2, no. 9, pp. 182-196. <https://doi.org/https://doi.org/10.26740/jrpipm.v9n2.p182-196>, 2025.
- [26] A. Faradilla, P. Wijayanti and R. Setianingsih, "Literasi Matematis Siswa SMP dalam Menyelesaikan Soal MATematika Model PISA dengan Konten Change and Relationship," *Pi: Mathematics Education Journal*, vol. 7, no. 2, pp. 72-83. <https://doi.org/10.21067/pmej.v7i2.10497>, 2024.
- [27] U. F. Handayani, "Berpikir Kreatif Siswa Kemampuan Matematika Tinggi dalam Memecahkan Masalah Kontekstual," *Pi: Mathematics Education Journal*, vol. 4, no. 2, p. 91–101. <https://doi.org/10.21067/pmej.v4i2.5757>, 2021.
- [28] T. Andriani, K. Suastika and N. R. Sesanti, "Analisis Kesalahan Konsep Matematika Siswa dalam Menyelesaikan Soal Trigonometri Kelas X TKJ SMKN 1 Gempol Tahun Pelajaran 2016/2017," *Pi: Mathematics Education Journal*, vol. 1, no. 1, p. 34–39. <https://doi.org/10.21067/pmej.v1i1.1998> , 2017.
- [29] L. Setyaningsih, "Pelaksanaan Lesson Study for Learning Community di MGMP Matematika SMP Kota Malang Tahun 2017," *Pi: Mathematics Education Journal*, vol. 1, no. 2, p. 47–56. <https://doi.org/10.21067/pmej.v1i2.2325>, 2018.
- [30] H. Lanya, M. Zayyadi, S. Aini and Haris, "Mathematic Communication Skills in Senior High School Based on Gender Differences," *Pi: Mathematics Education Journal*, vol. 3, no. 2, p. 75–81. <https://doi.org/10.21067/pmej.v3i2.4718>, 2020.