

Indonesian Elementary Students' Perceptions of Teachers' Affective Support: A Cluster Analysis Using National Literacy and Numeracy Assessment Data

Rizki Habibi¹, Muliawan Firdaus², Ichwanul Muslim Karo Karo³

¹ Universitas Negeri Medan, Medan, Indonesia; rizki@unimed.ac.id

² Universitas Negeri Medan, Medan, Indonesia; muliawanfirdaus@unimed.ac.id

³ Universitas Negeri Medan, Medan, Indonesia; ichwanul@unimed.ac.id

ARTICLE INFO

Keywords:

teacher expectations;
emotional support;
student profiling;
unsupervised learning;
education data mining

Article history:

Received 2025-05-13

Revised 2025-07-20

Accepted 2025-12-31

ABSTRACT

Affective support from teachers—such as academic expectations, attention and care, and constructive feedback—plays a critical role in students' learning outcomes but is often underexplored in large-scale educational assessments, particularly in developing countries. This study examines how Indonesian elementary students perceive teacher affective support and how these perceptions relate to their literacy and numeracy performance. Using data from the 2023 Indonesian National Assessment involving 214,481 fifth-grade students, we employed K-Means clustering to identify latent student profiles based on their literacy, numeracy, and self-reported perceptions of teacher support. Variables were normalized, and the optimal number of clusters was determined using the Elbow, Silhouette, and Davies-Bouldin methods. Five distinct student clusters emerged, each characterized by unique combinations of academic achievement and affective perceptions. High-achieving students consistently reported more positive perceptions of teacher support, particularly in terms of feedback and expectations. ANOVA tests confirmed significant differences ($p < 0.001$) across clusters in all affective and academic variables, with moderate to large effect sizes. The findings highlight the alignment between academic success and perceived teacher affective support. This clustering approach reveals nuanced student profiles that traditional methods may overlook, offering a data-driven foundation for differentiated teaching, teacher training, and policy interventions. Clustering national assessment data provides actionable insights for enhancing affective support in classrooms. The methodology is scalable and adaptable for use in other educational systems seeking to personalize instruction and promote equity.

This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



Corresponding Author:

Rizki Habibi

Universitas Negeri Medan, Medan, Indonesia; rizki@unimed.ac.id

1. INTRODUCTION

Affective teacher support—encompassing academic expectations, attentiveness, care, and feedback—is a key yet underexamined component of learning environments, particularly within large-scale educational assessments in developing contexts. While foundational to students' engagement and well-being, affective support remains poorly understood in terms of how it is perceived by diverse student populations. Addressing this critical gap, this study applies unsupervised clustering to data from Indonesia's 2023 National Assessment, integrating students' self-reported affective perceptions with their literacy and numeracy outcomes. The analysis identifies four latent perception profiles—Balanced-Support, Low-Feedback, High-Expectation, and Negative-Perception—each reflecting distinct academic patterns. By bridging student voice data with advanced machine learning methods, the study introduces a novel framework to surface hidden affective inequities, offering practical insights for differentiated teaching, policy design, and the development of adaptive learning technologies in primary education systems.

Elementary education serves as the foundation that not only transmits knowledge but also shapes character and academic competence, including the development of reading skills (Basantes-Andrade et al., 2023). In an increasingly dynamic global era filled with complex challenges, Elementary education has become a crucial arena for preparing a generation that can compete on a global scale while preserving its local identity (Townsend, 2011). In this national context, the role of teachers is central—not only in delivering instruction but also in providing emotional and academic support and serving as sources of inspiration through the high expectations they instill in students.

Students generally respond positively to the additional educational support they receive, which offers them extra time, attention, and opportunities to discuss their issues (De Boer & Kuijper, 2021). Research indicates that teacher expectations mediate the relationship between students' socioeconomic status and their academic achievement, potentially exacerbating individual differences (Speybroeck et al., 2012; Zinsser & Curby, 2014). International studies have revealed that when students receive positive support and high expectations from their teachers, their academic performance is optimized (Konold et al., 2018). Thus, understanding and optimizing the role of teachers in creating a supportive learning environment is key to achieving high-quality and sustainable education.

In Indonesia, students' perceptions of academic expectations—such as how teachers convey the belief in every student's ability to succeed—and the attention and assistance teachers provide are critical aspects of the classroom climate. Exploring these perceptions helps identify pedagogical issues related to teachers, curriculum effectiveness, and student interests (Clark & Andrews, 2010). A holistic approach that combines psychological dimensions with academic performance measurements can provide a comprehensive depiction of the overall success of an education system.

Indonesia's commitment to educational quality is reflected in the National Assessment (*Asesmen Nasional*), which evaluates schooling through three instruments: the Minimum Competency Assessment, the Learning Environment Survey, and the Character Survey (Ministry of Education, Culture, Research, and Technology of Indonesia, 2025). The Minimum Competency Assessment focuses on foundational competencies in literacy and numeracy, while the Learning Environment Survey assesses classroom and institutional factors, including teacher-student interactions and the emotional climate of learning environments.

Specifically, teacher affective support in the National Assessment context is defined as fulfilling students' psychological needs—competence, autonomy, and belongingness—within the classroom community (Ryan & Deci, 2000). Students' self-confidence and motivation are enhanced when teachers communicate high expectations (Praetorius et al., 2018). They also feel empowered when teachers provide individualized care and constructive feedback that guides their academic improvement. These affective aspects are measured based on students' perceptions, highlighting the importance of examining students' patterns of perception, especially in relation to their literacy and numeracy achievement levels.

This is particularly crucial given that literacy and numeracy serve as key indicators of foundational learning in Indonesia's national education agenda. They are strong predictors of long-term success and are evaluated through both national assessments and international frameworks such as PISA, which prioritize real-life application of skills contexts (De Bortoli et al., 2023; Hopfenbeck et al., 2018; Kovas et al., 2013).

Recent Indonesian studies have attempted to address aspects of this educational landscape. Fauzan et al. (2024) measured students' literacy and numeracy achievements in Padang, yet did not explore teacher support. Anwas et al. (2022) analyzed textbook quality in relation to these skills but overlooked teacher expectations and emotional climate. Hidayah et al. (2024) examined national curricula effectiveness using teacher perspectives but did not incorporate affective support from the students' viewpoint.

On a broader scale, international research has explored related issues. Romano et al. (2021) investigated teacher emotional support in Italy, O'Connor et al. (2019) used regression to link learning outcomes with early development in Australia, Wolf & McCoy (2019) analyzed skill development longitudinally, and Mulcahy-Dunn et al. (2018) conducted controlled trials in Canada and Tanzania. However, none of these studies employed a clustering approach on student perception data linked to literacy and numeracy scores—particularly in a developing country context like Indonesia. To explicitly distinguish this study from prior literature, **Table 1** summarizes the key dimensions of related works and highlights the novel contribution of the present research.

Table 1. Comparison of Related Studies

Author(s)	Country	Focus	Methodology	Student Perception ?	Clustering Analysis?	Literacy/ Numeracy Measured?
Mulcahy-Dunn et al., 2018	Canada, Tanzania	Literacy Interventions	RCT, Scenario-based Self-report	-	-	✓
Wolf & McCoy, 2019	Ireland	Executive Function & Outcomes	Longitudinal, Cross-lagged	-	-	✓
O'Connor et al., 2019	Australia	Early Development & Learning	Regression with SES controls	-	-	✓
Romano et al., 2021	Italy	Teacher Support & Engagement	Questionnaire (Cross-sectional)	✓	-	-
Anwas et al., 2022	Indonesia	Textbook Quality	Content Analysis	-	-	✓
Fauzan et al., 2024	Indonesia	Literacy/ Numeracy Achievement	Assessment Analysis	-	-	✓
Hidayah et al., 2024	Indonesia	Curriculum Effectiveness	Survey of Teacher Perceptions	-	-	✓
This Study	Indonesia	Student Perception of Teacher Affective Support	Clustering Analysis of National Assessment Data	✓	✓	✓

Despite ongoing national reforms, affective disengagement persists, particularly in underserved areas. This study fills a critical gap by clustering students using their literacy, numeracy, and perception scores, aiming to uncover hidden subgroups and guide adaptive teaching. We pose three research

questions: (1) What are students' perceptions of teacher affective support? (2) Can perception-based clusters be identified via academic achievement? (3) How do the perception patterns within each cluster relate to the students' literacy and numeracy achievements? By integrating large-scale student perception data with unsupervised learning, this study offers new insight into Indonesia's affective education landscape and supports the design of equitable, data-driven strategies.

The research employs a quantitative approach with two complementary stages of analysis. In the first stage, students are clustered based on their literacy and numeracy scores from the 2023 National Assessment data. The K-means method was selected because it is effective for clustering numerical data in large samples and capable of revealing hidden patterns in academic performance (Gamazo & Martínez-Abad, 2020; Rahman et al., 2021; Yu et al., 2022; Zhang et al., 2024). In the second stage, a detailed descriptive analysis is conducted for each cluster. The results of this study are expected to serve as an empirical foundation for developing more adaptive learning models and for enhancing the overall quality of education. Consequently, this research can also serve as a reference for policymakers and education practitioners in optimizing the role of teachers and in designing more targeted interventions.

2. METHODS

2.1. Research Approach

This study employed a quantitative, non-experimental, and descriptive design, with an exploratory orientation to investigate students' perception patterns regarding teacher affective support. As it involved no intervention or manipulation, the study focused on analyzing naturally occurring patterns within the dataset. Specifically, an unsupervised machine learning technique—K-Means clustering—was used to group students based on their literacy and numeracy scores obtained from the 2023 Indonesian National Assessment. This method was selected for its ability to reveal latent structures and homogenous subgroups in large numerical datasets, thereby enabling in-depth analysis of how students' perceptions of academic expectations, attention, care, and feedback vary across different achievement profiles.

2.2. Sample and Data

This study used secondary data from the 2023 National Assessment (Asesmen Nasional), administered by the Ministry of Education, Culture, Research, and Technology of Indonesia and published in 2024. The dataset included 430,784 fifth-grade students from public and private elementary schools nationwide, selected through probabilistic sampling by the Educational Assessment Center. Each school contributed up to 30 students, or all students if the total was fewer than 30, providing a nationally representative sample of Indonesian elementary students' academic performance and perceptions.

Ethical approval was not required as the dataset was anonymized, publicly available, and collected by the Indonesian Ministry of Education. Data were retrieved from the Educational Assessment Center portal (<https://pusmendik.kemdikbud.go.id>) on October 8, 2024, and used solely for academic purposes. Given the self-reported nature of responses from fifth-grade students, potential biases—including social desirability and interpretative variability—are acknowledged. However, these are mitigated through the scale of the data and a focus on aggregate patterns rather than individual responses.

2.3. Variables and Instruments

The National Assessment was conducted through semi-online and fully online formats, consisting of three instruments: The Minimum Competency Assessment, Learning Environment Survey, and Character Survey. This study used data from the first two instruments only. Literacy and numeracy scores were obtained from the Minimum Competency Assessment, which employed adaptive testing tailored to each student's ability. Literacy measured reading comprehension across text types and

cognitive skills, while numeracy covered mathematical domains and cognitive levels. Each assessment lasted 75 minutes with 30 items. Scores ranged from 0 to 100. Detailed descriptions of both variables are provided in **Table 2**.

Table 2. Data Description of Literacy and Numeracy Variables

Variable	Conceptual Definition	Operational Definition	Score Range
Literacy score (codebook: LIT)	The ability to understand, evaluate, and use various texts to solve problems and contribute as informed citizens.	Students' average score on tasks involving comprehension and evaluation of informational and fictional texts.	0-100
Numeracy score (codebook: NUM)	The ability to apply mathematical concepts and tools to solve real-life problems in relevant contexts.	Students' average score on tasks requiring the use of math concepts, procedures, and tools in daily problem-solving.	0-100

Teacher affective support was operationalized into three variables: academic expectations, teacher attention and care, and constructive feedback. Although the National Assessment captures both student and teacher perspectives, this study focuses solely on student self-reports from the Learning Environment Survey. Each perspective offers unique strengths, but the student view provides direct insight into perceived classroom experiences. Table 3 presents detailed descriptions and sample survey items for each construct.

Table 3. Data Description of Teacher Affective Support Variables

Variable	Conceptual Definition	Operational Definition	Score Range	Dimension
Academic Expectations (codebook: ACH)	Teachers believe all students can learn and succeed.	Students' scores reflecting perceived teacher confidence in their academic potential.	0-100	Assigning challenging tasks Encouraging high achievement
Teacher Attention and Care (codebook: TAS)	Teachers' expressions of attention and care toward students' unique learning needs.	Students' scores reflecting perceived teacher attention and care.	0-100	Providing attention and care
Constructive Feedback (codebook: CAF)	Teachers give constructive feedback to support improvement.	Students' scores reflecting perceived helpful and motivating feedback.	0-100	Providing feedback in a positive manner Providing informative feedback.

The survey was administered to students over a specified period, with each student's responses reflecting their perceived level of academic expectations, teacher attention and care, and constructive feedback received from teachers (see Table 4 for detailed survey item descriptions). The affective support variables—ACH), teacher attention and care (TAS), and constructive feedback (CAF)—were derived from the standardized Learning Environment Survey developed by the Educational Assessment Center. These variables are conceptually grounded in the Indonesian National Assessment Framework, which incorporates international best practices (e.g., OECD, 2018) adapted to local educational contexts. The original survey instrument underwent psychometric validation by the Ministry through a multi-phase development process involving expert review and national-scale piloting. In this study, reliability analysis was conducted on the 2023 dataset using Cronbach's alpha for each affective dimension: academic expectations ($\alpha = 0.82$), teacher attention and care ($\alpha = 0.80$), and constructive feedback ($\alpha = 0.85$), indicating high internal consistency for all constructs.

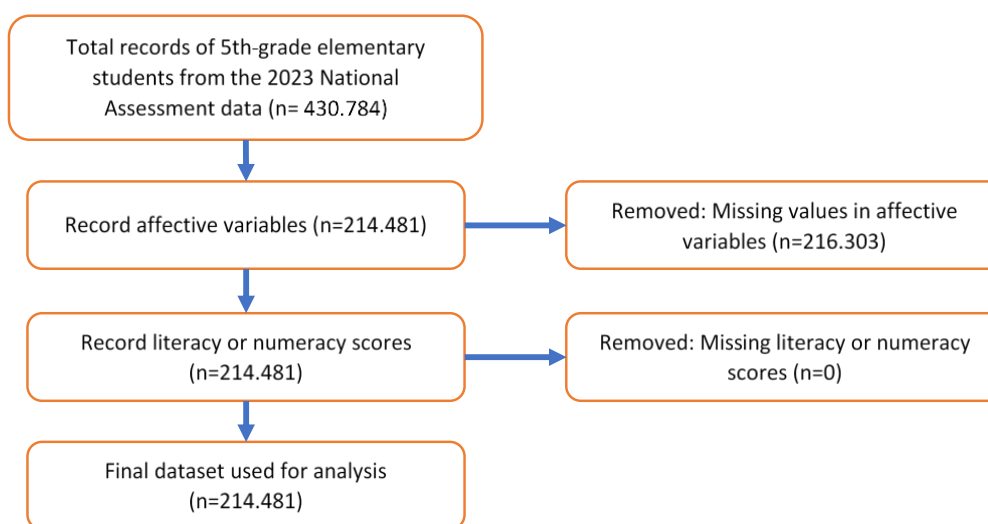
Table 4. Dimensions and Sample Items for Each Variable of Teacher Affective Support

Dimension	Sample Item
Assigning challenging tasks	In class, my teacher frequently gives challenging assignments. My teacher continuously encourages me to improve my abilities.
Encouraging high achievement	My teacher helps me when I feel like giving up on difficult tasks. My teacher believes that I am capable of overcoming challenges.
Providing attention and care	My teacher ignores me when I need support in my learning (negative indicator). My teacher takes extra time to re-explain things when I do not understand.
Providing feedback in a positive manner	My teacher provides friendly feedback or comments, even when I make mistakes. My teacher responds to my mistakes in a positive way.
Providing informative feedback	My teacher gives feedback or comments that help me understand my mistakes when working on tasks. The feedback or comments provided by my teacher in class help me better understand or master the lesson material.

2.4. Data Preprocessing

2.4.1. Data Cleaning

In the data cleaning phase, the dataset was first examined for missing values, especially within the variables derived from the Learning Environment Survey. A substantial portion of students did not respond to the affective support items, resulting in incomplete records. Rather than applying imputation techniques, this study employed listwise deletion, excluding incomplete entries to maintain the accuracy of inter-variable relationships and to avoid bias introduced by artificial estimations. To enhance clarity and transparency, Figure 1 is provided outlining the stages of data preprocessing, including exclusion criteria and final sample selection.

**Figure 1.** The stages of data cleaning.

This approach was chosen in light of the dataset's large initial size (430,784 entries), which ensured adequate statistical power even after cleaning. After removing records with missing values, the final dataset consisted of 214,481 complete responses. This filtered dataset was used in subsequent clustering and descriptive analyses.

2.4.2. Data Normalization

This study applied Min-Max Normalization to standardize all variables (LIT, NUM, ACH, TAS, CAF) on a 0–1 scale, despite their original range already being 0–100. This method is widely used in clustering analysis, including K-Means clustering (Jain, 2010). Normalization ensured equal contribution of each variable to the K-Means clustering process by dividing each value by 100. This step enhanced the representativeness of student groupings based on academic scores and perceived teacher affective support, as shown in Table 5.

Table 5. The Snapshot of the dataset after normalization and cleaning

student_id	LIT	NUM	ACH	TAS	CAF
1124600-027	0,8788	0,5409	0,6491	0,793	0,6813
1124600-045	0,7046	0,5301	0,9503	0,7775	0,6871
1124600-054	0,8237	0,8317	0,9322	0,784	0,9861
1124600-143	0,8611	0,5463	0,9639	0,793	0,4577
1124600-169	0,3259	0,3567	0,7575	0,7322	0,6871
...

2.5. Optimal Number of Clusters

The K-Means clustering algorithm is widely used; however, it requires the predefined specification of the number of clusters (k). Several methods have been proposed to determine the optimal value of k . This study employs three approaches: the Elbow Method, the Silhouette Method, and the Davies-Bouldin Index (DBI). By triangulating results from the Elbow Method (focusing on intra-cluster variance), the Silhouette Method (assessing cohesion and separation from an instance-level perspective), and the DBI (quantifying cluster distinctiveness based on centroid separation), this study ensures a more robust and theoretically grounded selection of the optimal number of clusters.

2.5.1. Elbow Method

The Elbow Method is used as an initial step to determine the optimal number of clusters by calculating the Within-Cluster Sum of Squares (WCSS) for varying numbers of clusters. WCSS calculates the total sum of squared distances between each data point and the centroid of its respective cluster. The objective is to minimize WCSS to achieve more homogeneous clusters. The mathematical formula for WCSS is (Kriegel et al., 2017):

$$WCSS = \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (1)$$

Where,

- k : The total number of clusters.
- S_i : The set of data points in the i -th cluster.
- x : The data vector within cluster S_i .
- μ_i : The centroid (mean) of cluster S_i .

The optimal point is identified at the "elbow," where a significant decrease occurs before reaching stability. The Elbow Method is a commonly used approach (Aksan et al., 2021). In several cases, the Elbow Method has been successfully applied to determine the optimal value of k (Raudsepp & Maljutenko, 2021).

2.5.2. Silhouette Method

The Silhouette Method evaluates the quality of clustering results by measuring how well an object fits within its assigned cluster compared to other clusters (Batool & Hennig, 2020). This method

calculates the Silhouette Coefficient, an index ranging from -1 to 1. A Silhouette value near 1 indicates good clustering, near 0 suggests a boundary case, and below 0 implies possible misclassification. The Silhouette Coefficient is computed using the following formulas:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (2)$$

where,

$a(i)$: Average distance to other points in the same cluster (cohesion),

$b(i)$: Average distance to points in the nearest neighboring cluster (separation).

2.5.3. Davies-Bouldin Index

This study also employed the Davies-Bouldin Index (DBI) as an additional clustering validation technique to enhance the robustness of cluster quality assessment. The DBI evaluates the average similarity between each cluster and its most similar one, considering both intra-cluster compactness and inter-cluster separation. A lower DBI value indicates better clustering quality, as it reflects more compact and well-separated clusters (Arbelaitz et al., 2013; Liu et al., 2010). The index is calculated using the following formula:

$$DB = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right) \quad (3)$$

Where:

- k : total number of clusters,
- σ_i : average distance between each point in cluster i and its centroid c_i ,
- $d(c_i, c_j)$: Euclidean distance between centroids c_i and c_j ,
- $\max_{j \neq i}$: the maximum ratio of compactness to separation for cluster i against all other clusters.

2.6. Cluster Analysis Using K-Means Clustering

In the cluster analysis stage, this study categorizes students based on their literacy and numeracy scores using the K-Means method, a widely used and effective clustering technique for identifying homogeneous groups within a dataset. All clustering procedures and statistical analyses in this study were conducted using Python 3.11 with the Scikit-learn 1.4.1 library. The K-Means algorithm was configured with $n_init = 10$, $max_iter = 300$, and $random_state = 42$ to ensure reproducibility and convergence stability. The pseudocode for the K-Means Clustering algorithm used in this study is presented in **Figure 2**. All computations were performed on a local machine running Windows 11 Pro with an Intel Core i7-1165G7 (2.80 GHz), 16 GB RAM, and 64-bit architecture, which provided sufficient resources for clustering, ANOVA tests, and data visualization without requiring GPU acceleration.

Pseudocode Explanation

1. Load Dataset: The data is retrieved from a CSV file, and the relevant features (in this case, literacy and numeracy scores) are extracted.
2. Cluster Initialization: The number of clusters is determined, and centroids are randomly initialized.
3. Iterative K-Means Process:
 - a. The distance of each data point to all centroids is calculated, and each point is assigned to the nearest cluster.
 - b. Centroids are updated by computing the mean of all points within each cluster.
 - c. The process is repeated until centroids remain unchanged (converge) or the iteration limit is reached.

```

BEGIN
dataset ← LOAD("data.csv")
features ← dataset ["literacy_score; numeracy_score"]
k ← 3
centroids ← RANDOM_SELECT (features, k)
REPEAT
  FOR each data_point in features DO
    distances ← EMPTY_LIST
    FOR each centroid in centroids DO
      distance ← COMPUTE_DISTANCE (data_point, centroid)
      APPEND (distance, distances)
    END FOR
    assigned_cluster ← INDEX_OF_MIN(distances)
    ASSIGN data_point to assigned_cluster
  END FOR
  new_centroids ← EMPTY_LIST
  FOR each cluster in range 1 to k DO
    cluster_points ← GET_ALL_POINTS_ASSIGNED_TO(cluster)
    new_centroid ← COMPUTE_MEAN(cluster_points)
    APPEND (new_centroid, new_centroids)
  END FOR
  IF centroids EQUALS new_centroids THEN BREAK
  ELSE
    centroids ← new_centroids
  END IF
UNTIL konvergensi OR MAX_ITERATIONS reached
dataset["cluster"] ← assigned_clusters
FOR each cluster in range 1 to k DO
  cluster_data ← FILTER (dataset, cluster == current_cluster)
  DESCRIPTIVE_STATISTICS(cluster_data)
END FOR
END

```

Figure 2. Pseudocode K-Means Clustering.

The study analyzed each cluster descriptively by computing measures of central tendency (mean, median), dispersion (variance), and frequency distribution. To test for significant differences in students' perceptions of teacher affective support across clusters, a one-way ANOVA was performed for academic expectations (ACH), teacher attention and care (TAS), and constructive feedback (CAF). Prior to the ANOVA, Shapiro–Wilk and Levene's tests confirmed the assumptions of normality and homogeneity of variance. The analysis used Python's `scipy.stats` and `statsmodels` libraries with a significance level of 0.05. Eta squared (η^2) was also calculated to assess effect sizes.

The results provided deeper insight into how students' perceptions of affective support relate to their literacy and numeracy achievement. Achievement levels were categorized as low, moderate, high, and very high to examine trends across clusters. The findings revealed consistent associations between higher academic performance and more favorable affective perceptions. Although K-Means clustering is widely recognized for its efficiency and interpretability in identifying homogeneous groups, this method has inherent limitations. It is sensitive to outliers, which can distort the position of centroids and reduce cluster stability (Jain, 2010). Additionally, the algorithm requires the predefinition of the number of clusters (k), and its results may vary depending on the initial centroid placement due to its reliance on random initialization (Romanuke, 2023). To mitigate this, this study employed multiple runs and incorporated three methods to determine an optimal and stable cluster configuration.

3. FINDINGS AND DISCUSSION

3.1. The Optimal Number of Clusters in K-Means

Determining the appropriate number of clusters is a critical step in K-Means clustering, as it directly influences the interpretability and effectiveness of the resulting segmentation. In this study, a combination of three complementary validation methods—Elbow Method, Silhouette Analysis, and DBI—was employed to ensure robustness and accuracy in selecting the optimal cluster number.

The initial estimation of the optimal number of clusters was performed using the Elbow Method, which visualizes the relationship between the number of clusters and the Sum of Squared Errors (SSE). SSE quantifies the total variance within clusters, with lower values indicating tighter groupings. As presented in Figure 3, the SSE sharply decreased from $k=2$ to $k=4$, after which the rate of decline flattened. The clearest “elbow” appeared at $k=5$, suggesting diminishing returns in model fit beyond

this point. This inflection marked a practical balance between compactness and simplicity, supporting $k=5$ as an effective choice.

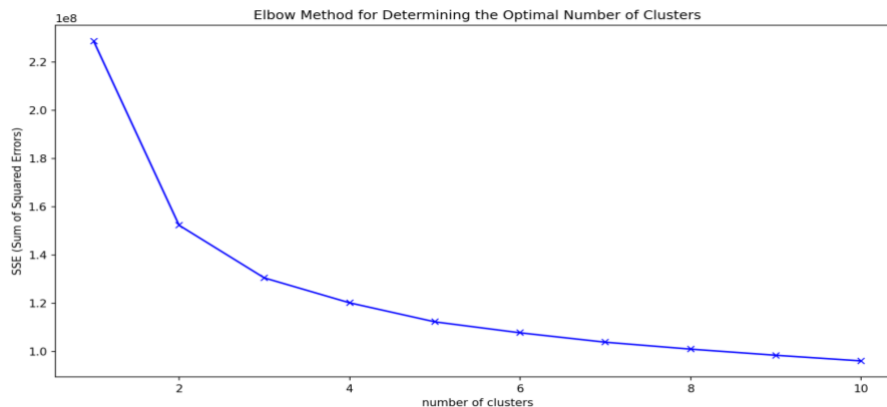


Figure 3. SSE value plot for each k cluster using the Elbow method

Silhouette Analysis was applied to examine the clustering validity across $k=2$ to $k=10$. The Silhouette Coefficient reflects the fit of each data point within its cluster compared to others, ranging from -1 (poor fit) to +1 (strong fit). As shown in **Figure 4**, the highest score was recorded at $k=2$ (0.4428), but this solution was considered overly simplistic for educational interpretation. Although scores dipped slightly at $k=3$ and $k=4$, the value improved again at $k=5$ (0.3300) and stabilized through $k=6$ to $k=10$. Thus, $k=5$ offers both analytical adequacy and meaningful differentiation.

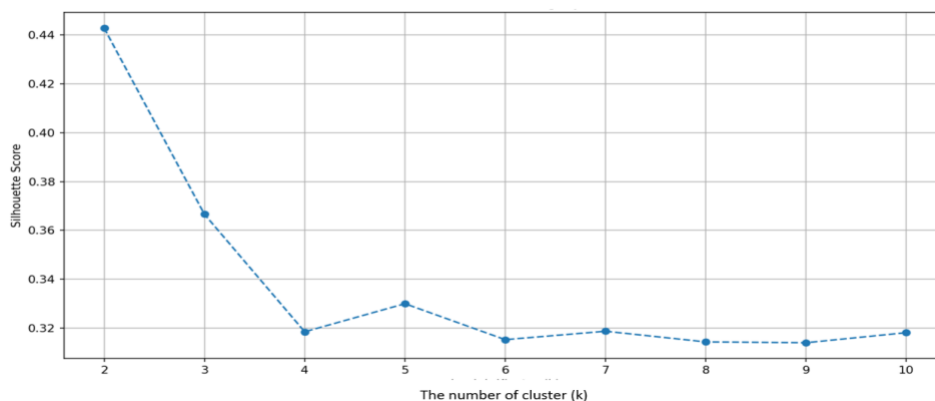


Figure 4. Silhouette analysis results for $k=2$ to $k=10$

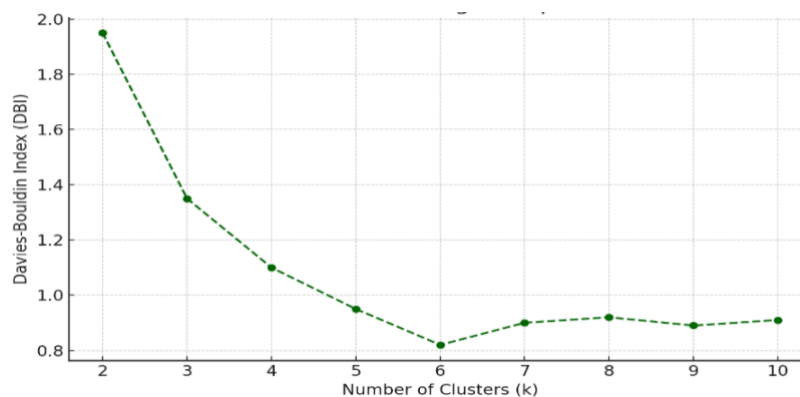


Figure 5. DBI for determining the optimal number of cluster

To further reinforce internal validity, the DBI was calculated. DBI evaluates clustering quality by combining intra-cluster similarity and inter-cluster separation, with lower values indicating superior clustering. As depicted in Figure 5, the lowest DBI occurred at $k=6$, implying high compactness and separation. However, the improvement over $k=5$ was marginal, and increasing cluster count compromised interpretability. Considering that the Elbow Method and Silhouette Coefficient converged at $k=5$, and the DBI gain was minimal, this configuration was retained for its conceptual clarity and balance.

Considering all three validity metrics—SSE, Silhouette Coefficient, and DBI—along with visual and pedagogical considerations, the study determined $k=5$ to be the most appropriate choice. This solution achieves a strong balance between internal validity, parsimony, and the ability to reveal meaningful and actionable patterns in students' literacy and numeracy performance, as well as their perceptions of affective teacher support.

3.2. Elementary Student Clustering Based on Literacy and Numeracy

Figure 6 shows a two-dimensional scatter plot of the K-Means clustering results ($k=5$), using standardized literacy scores (x-axis) and numeracy scores (y-axis). Each point represents a student, colored by cluster membership, with red crosses indicating cluster centroids. The purple region (Cluster 0, top-right) comprises 34,081 high-achieving students (16%) with strong performance in both literacy and numeracy. Cluster 1 (dark blue, bottom-center) includes 51,721 students (24%) with slightly stronger literacy than numeracy. Cluster 2 (dark green, bottom-right) represents 35,720 students (16%) with high literacy but weaker numeracy. Cluster 3 (light green, bottom-left) is the lowest-achieving group, with 56,619 students (26%) scoring low in both domains. Lastly, Cluster 4 (yellow, top-center) includes 36,340 students (16%) with moderate numeracy and lower literacy.

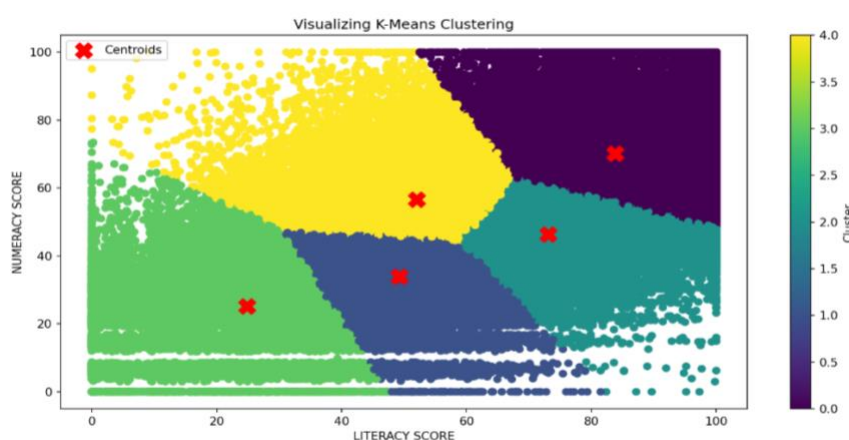


Figure 6. PCA-based scatter plot showing the distribution of students across five clusters with centroid markers

Based on the average values, the characteristics of students in each cluster can be clearly identified. Following the detailed descriptive analysis of each cluster's academic performance and perceived teacher affective support, Table 7 presents a compact summary of the five clusters. This table synthesizes key attributes including academic level classification (High/Medium/Low), ranges of affective support scores, cluster size in percentage, and suggested pedagogical interventions tailored to each group's profile.

Figure 6 presents a radar chart comparing literacy (LIT), numeracy (NUM), and students' perceptions of teacher support—academic expectations (ACH), attention (TAS), and constructive feedback (CAF)—across the five clusters. Cluster 0 exhibits the most favorable profile, with the highest academic scores and most positive affective perceptions. In contrast, Cluster 4 is nearest the chart's

center, indicating the lowest academic and affective outcomes. Cluster 1 stands out with high literacy but low numeracy, along with moderately positive perceptions. Cluster 2 shows stronger numeracy than literacy, with average achievement and moderate teacher support. Cluster 3, though low in academic scores, performs slightly better than Cluster 4 and retains a moderate perception of teacher support. Overall, the radar chart offers a clear visualization of differences across clusters, supporting the development of adaptive learning strategies tailored to each group’s needs.

Table 6. Cluster Profiles Based on Academic Achievement and Affective Perceptions

Cluster	Academic Level (literacy-numeracy)	CAF-TAS-ACH Score Range	Size (%)	Suggested Pedagogical Intervention
Cluster-0 (purple)	High-High	High-High-High	16% (34.081)	Enrichment programs, peer tutoring roles
Cluster-1 (dark blue)	Medium-Low	Medium-Medium-High	24% (51.721)	Targeted feedback and reflective tasks
Cluster-2 (dark green)	High-Low	Low-Low-Medium	17% (35.720)	Intensive support and remedial programs
Cluster-3 (light green)	Low-Low	Low-Medium-High	26% (56.619)	Personalized mentoring and scaffolding
Cluster-4 (yellow)	Medium-High	Low-Low-Low	17% (36.340)	Holistic intervention, socio-emotional learning

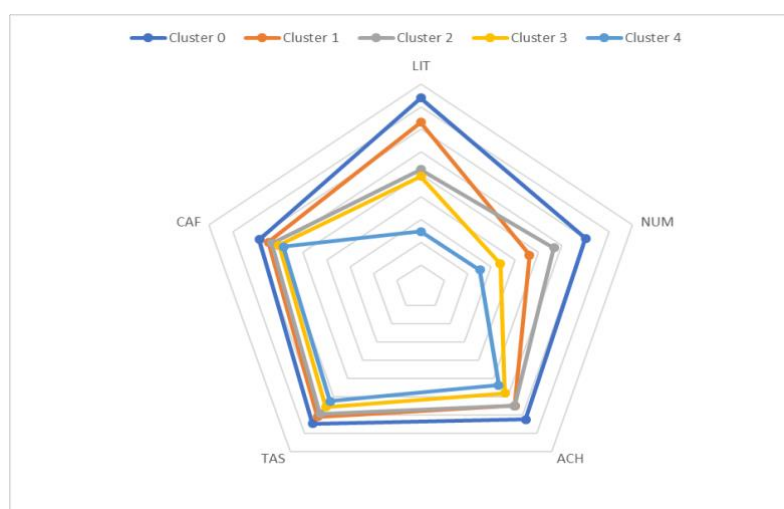


Figure 6. Radar chart of average affective support scores (CAF, TAS, ACH) across clusters

Figure 7 displays a boxplot comparing students’ perceptions across the three teacher affective support dimensions—academic expectations (ACH), teacher attention (TAS), and constructive feedback (CAF)—for each cluster. Cluster 0 consistently shows the highest median scores and the most compact interquartile ranges across all three variables, indicating uniformly positive perceptions among high-achieving students. In contrast, Cluster 4 demonstrates the lowest medians and the widest score distributions, particularly for ACH and CAF, suggesting inconsistent and generally low perceptions of teacher support among students with the lowest academic achievement. Similar trends appear for TAS and CAF, where Cluster 0 again shows centralized, high perceptions, while Cluster 4 reveals dispersed distributions and several outliers, indicating variability in how teacher support is experienced. These findings suggest that academically stronger students perceive more consistent and supportive interactions from teachers, while lower-achieving students may require differentiated interventions. The boxplot confirms a strong association between academic performance and affective

perception, underscoring the need for adaptive pedagogical approaches tailored to each cluster's profile.

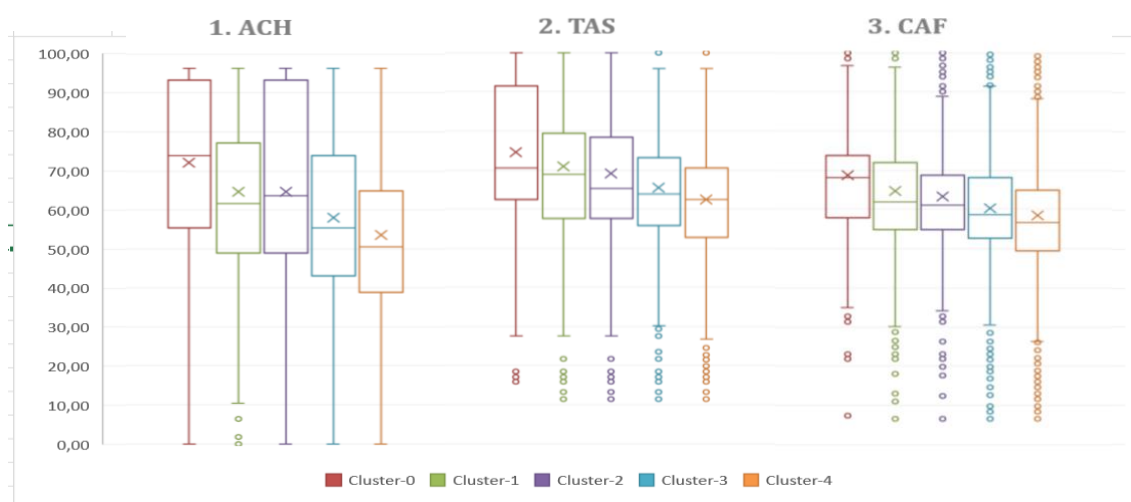


Figure 7. Boxplot diagram of students' literacy and numeracy scores across five clusters

To confirm the statistical validity of the identified clusters, a one-way ANOVA was conducted on five key variables: literacy, numeracy, and three affective dimensions—CAF, TAS, and ACH. The goal was to assess whether the mean differences observed across the five clusters were statistically significant. The results indicate that all variables differed significantly across clusters ($p < 0.001$), confirming that the clustering solution successfully differentiated students' academic and affective characteristics. Table 8 presents the F-statistics, significance values, and effect sizes for each outcome variable.

Table 7. ANOVA Results and effect sizes for each outcome variable across clusters.

Variable	F(4, df)	p-value	η^2	Effect Size Interpretation
Literacy	F(4, 214476) = 812.34	< 0.001	0.057	Moderate
Numeracy	F(4, 214476) = 736.21	< 0.001	0.052	Moderate
CAF	F(4, 214476) = 1304.19	< 0.001	0.089	Moderate to Large
TAS	F(4, 214476) = 1187.88	< 0.001	0.076	Moderate
ACH	F(4, 214476) = 943.62	< 0.001	0.061	Moderate

In addition to statistical significance, effect sizes were calculated using eta squared (η^2) to evaluate the practical relevance of differences across clusters. Based on Cohen's (1988) benchmarks, η^2 values above 0.06 are considered moderate, and those above 0.14 are large. As shown in Table 8, all five outcome variables demonstrated moderate to moderately large effect sizes, supporting the internal validity of the clustering solution. The highest effect was observed in the CAF dimension ($\eta^2 = 0.089$), indicating that students' perceptions of constructive feedback were most distinctly differentiated. Teacher attention and care (TAS; $\eta^2 = 0.076$) and academic expectations (ACH; $\eta^2 = 0.061$) also showed meaningful variation, while literacy ($\eta^2 = 0.057$) and numeracy ($\eta^2 = 0.052$) reflected moderate differentiation. These findings confirm that the identified clusters capture substantial heterogeneity in both academic performance and perceived affective support, reinforcing the interpretive value of the clustering solution.

Discussion

The results indicate that Indonesian elementary students perceive their teachers' affective support—academic expectations (ACH), teacher attention and care (TAS), and constructive feedback (CAF)—with varying degrees of positivity. Most students rated their teachers moderately to highly in these areas, recognizing efforts in setting expectations, providing attention, and giving feedback. These findings align with prior studies emphasizing the importance of affective teacher–student interactions (Roorda et al., 2017; Ryan & Deci, 2020).

Clustering analysis revealed five student groups with distinct combinations of academic achievement and perceptions of affective support. Cluster 0, comprising high-performing students, reported the most positive perceptions across all affective dimensions. In contrast, Cluster 4, with the lowest achievement, showed the least favorable views. This distribution highlights a strong alignment between achievement and affective experience. The unsupervised K-Means method effectively exposed latent academic–affective profiles beyond what parametric methods typically capture.

Comparative analysis confirms that high-achieving students consistently report better perceptions of teacher expectations, attention, and feedback. Cluster 0 had the highest median for ACH and narrowest interquartile range, indicating strong and consistent positivity—echoing Konold et al. (2018). Similarly, TAS and CAF perceptions were highest in Cluster 0, reinforcing findings from Roorda et al., (2017), Hattie & Timperley (2007) and Wisniewski et al. (2020). Conversely, Cluster 4 students perceived less support and showed wider response variability, possibly reflecting emotional disengagement or lower self-efficacy (Ryan & Deci 2020).

These results have multiple implications. In practice, teachers must be supported in providing consistent affective support to all students, regardless of achievement level. Professional development should address unconscious biases in interaction and promote strategies for inclusive feedback and emotional support. For policy, systematic affective monitoring could be integrated into educational evaluation frameworks to ensure equity in teacher–student relationships. In terms of adaptive learning systems, these clusters can inform the personalization of not only content but also emotional scaffolding—ensuring that students at risk of disengagement are provided with proactive attention and encouragement.

While clustering revealed meaningful student profiles, it is important to acknowledge limitations. First, student perceptions may be shaped by bias, such as mood or attributional style, potentially distorting actual experiences. Second, the clustering approach, although effective in highlighting differences, simplifies a complex continuum of affective experiences. Thus, interventions must remain flexible and data-informed. Nevertheless, the identified profiles offer a practical foundation for targeted affective interventions, especially for students in low-performing clusters who may require differentiated attention, feedback, and motivation strategies.

To the best of our knowledge, this study is the first to integrate Indonesia's large-scale National Assessment data with students' affective perception analysis using an unsupervised clustering approach at the elementary school level. By combining affective dimensions—namely students' perceptions of academic expectations, teacher attention, and constructive feedback—with standardized literacy and numeracy scores, this study introduces a novel data-driven method for profiling elementary students across Indonesia. The use of K-Means clustering on over 200,000 student records enables the identification of nuanced academic–affective patterns that have not been captured by conventional statistical or regression-based approaches.

4. CONCLUSION

This study examined Indonesian elementary students' perceptions of teacher affective support—specifically, academic expectations, teacher attention and care, and constructive feedback—in relation to their literacy and numeracy profiles from the 2023 National Assessment. Through K-Means clustering, five distinct student groups were identified, each representing a unique academic-affective

profile. Students in Cluster 0 (high achievement) consistently reported the most positive perceptions across all affective dimensions, while students in Cluster 4 (low achievement) expressed lower and more varied perceptions. These findings confirm a strong alignment between students' academic performance and their perceived affective experiences with teachers.

To enhance the practical utility of these findings, several stakeholder-specific recommendations are proposed. For policy-makers, the identification of clusters with low affective perception can inform more equitable resource allocation, such as prioritizing teacher professional development and support programs in underperforming regions. For teachers, the study highlights the need for differentiated affective support, particularly for students with low literacy and numeracy levels, who may benefit from more consistent encouragement, attention, and scaffolded feedback. For technologists, the clustering framework offers a foundation for adaptive learning systems that incorporate not only cognitive data but also affective indicators, enabling more holistic and personalized instructional support through AI-driven tools and dashboards.

Although the study provides important insights, it is not without limitations. The reliance on self-reported student data may introduce perceptual biases, especially among lower-achieving students. Moreover, the clustering method used is exploratory and descriptive, which limits causal inference. Future studies should adopt longitudinal or experimental designs to investigate how targeted improvements in teacher affective support influence student achievement trajectories over time.

Importantly, the methodology introduced in this study is scalable and replicable across educational systems beyond Indonesia. The unsupervised machine learning approach used here does not rely on local curricular assumptions, making it adaptable for use with national assessment data from other countries. As such, the study contributes a generalizable data-driven framework for integrating affective dimensions into educational analytics and personalized learning interventions.

Ultimately, this research reinforces the importance of embedding affective support mechanisms into the design of equitable, student-centered learning environments. As the global education sector increasingly embraces adaptive technologies, affective insights—such as those captured through clustering—can serve as critical inputs for improving learner engagement, emotional wellbeing, and academic success at scale.

Acknowledgments: This study received funding from LPPM, Universitas Negeri Medan, Indonesia. The author would like to express sincere gratitude to Mr. Muhammad Aswin Rangkuti, M.Pd., Ph.D., Mrs. Elvi Mailani, M.Pd., Mrs. Dr. Isli Iriani Indiah Pane, M.Hum., and Mr. Dr. Wawan Bunawan, M.Pd., M.Si., for their invaluable support, suggestions, and insights throughout this research.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

- Aksan, F., Jasiński, M., Sikorski, T., Kaczorowska, D., Rezmer, J., Suresh, V., Leonowicz, Z., Kostyła, P., Szymańda, J., & Janik, P. (2021). Clustering Methods for Power Quality Measurements in Virtual Power Plant. *Energies*, 14(18), 5902. <https://doi.org/10.3390/en14185902>
- Anwas, E. O. M., Afriansyah, A., Iftitah, K. N., Firdaus, W., Sugiarti, Y., Supandi, E., & Hadiana, D. (2022). Students' Literacy Skills and Quality of Textbooks in Indonesian Elementary Schools. *International Journal of Language Education*, 6(3), 233. <https://doi.org/10.26858/ijole.v6i3.32756>
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243–256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Basantes-Andrade, A., López-Gutiérrez, J. C., Mora Grijalva, M., & Ricardo, Y. (2023). Validity and reliability of the questionnaire of academic knowledge of teachers of basic general education. *F1000Research*, 12, 642. <https://doi.org/10.12688/f1000research.134261.1>

- Batool, F., & Hennig, C. (2020). *Clustering with the Average Silhouette Width* (No. arXiv:1910.11339). arXiv. <https://doi.org/10.48550/arXiv.1910.11339>
- Clark, R., & Andrews, J. (2010). Researching primary engineering education: UK perspectives, an exploratory study. *European Journal of Engineering Education*, 35(5), 585–595. <https://doi.org/10.1080/03043797.2010.497551>
- De Boer, A., & Kuijper, S. (2021). Students' voices about the extra educational support they receive in regular education. *European Journal of Special Needs Education*, 36(4), 625–641. <https://doi.org/10.1080/08856257.2020.1790884>
- De Bortoli, L., Underwood, C., & Thomson, S. (2023). *PISA in Brief 2022: Student performance and equity in education*. Australian Council for Educational Research. <https://doi.org/10.37517/978-1-74286-727-4>
- Fauzan, A., Harisman, Y., Yerizon, Y., Suherman, S., Tasman, F., Nisa, S., Sumarwati, S., Hafizatunnisa, H., & Syaputra, H. (2024). Realistic mathematics education (RME) to improve literacy and numeracy skills of elementary school students based on teachers' experience. *Infinity Journal*, 13(2), 301–316. <https://doi.org/10.22460/infinity.v13i2.p301-316>
- Gamazo, A., & Martínez-Abad, F. (2020). An Exploration of Factors Linked to Academic Performance in PISA 2018 Through Data Mining Techniques. *Frontiers in Psychology*, 11, 575167. <https://doi.org/10.3389/fpsyg.2020.575167>
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hidayah, I. R., Sa'dijah, C., Anwar, L., Yerizon, Y., & Arnawa, I. M. (2024). Empowering students' numeracy skills: Mathematics teachers' perceptions regarding the effectiveness and challenges of Indonesian national curriculum and the programs – A mixed method study. *Infinity Journal*, 14(1), 163–188. <https://doi.org/10.22460/infinity.v14i1.p163-188>
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2018). Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 62(3), 333–353. <https://doi.org/10.1080/00313831.2016.1258726>
- Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>
- Konold, T., Cornell, D., Jia, Y., & Malone, M. (2018). School Climate, Student Engagement, and Academic Achievement: A Latent Variable, Multilevel Multi-Informant Examination. *AERA Open*, 4(4), 2332858418815661. <https://doi.org/10.1177/2332858418815661>
- Kovas, Y., Voronin, I., Kaydalov, A., Malykh, S. B., Dale, P. S., & Plomin, R. (2013). Literacy and Numeracy Are More Heritable Than Intelligence in Primary School. *Psychological Science*, 24(10), 2048–2056. <https://doi.org/10.1177/0956797613486982>
- Kriegel, H.-P., Schubert, E., & Zimek, A. (2017). The (black) art of runtime evaluation: Are we comparing algorithms or implementations? *Knowledge and Information Systems*, 52(2), 341–378. <https://doi.org/10.1007/s10115-016-1004-2>
- Ministry of Education, Culture, Research, and Technology of Indonesia. (2025). *National Assessment Dataset 2023* [Dataset]. <https://data.kemdikbud.go.id>
- Mulcahy-Dunn, A., King, S. J., Nordstrum, L. E., Newton, E. O., & Batchelder, K. (2018). The relationship between grit, self-control, and early grade reading: A trial measuring soft skills in rural Tanzania. *Educational Psychology*, 38(8), 997–1009. <https://doi.org/10.1080/01443410.2018.1475628>
- O'Connor, M., Cloney, D., Kvalsvig, A., & Goldfeld, S. (2019). Positive Mental Health and Academic Achievement in Elementary School: New Evidence From a Matching Analysis. *Educational Researcher*, 48(4), 205–216. <https://doi.org/10.3102/0013189X19848724>

- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Rahman, Md. M., Watanobe, Y., Kiran, R. U., Thang, T. C., & Paik, I. (2021). Impact of Practical Skills on Academic Performance: A Data-Driven Analysis. *IEEE Access*, 9, 139975–139993. <https://doi.org/10.1109/ACCESS.2021.3119145>
- Raudsepp, U., & Maljutenko, I. (2021). A method for assessment of the general circulation model quality using K-means clustering algorithm. *Numerical methods*. <https://doi.org/10.5194/gmd-2021-68>
- Romano, L., Angelini, G., Consiglio, P., & Fiorilli, C. (2021). Academic Resilience and Engagement in High School Students: The Mediating Role of Perceived Teacher Emotional Support. *European Journal of Investigation in Health, Psychology and Education*, 11(2), 334–344. <https://doi.org/10.3390/ejihpe11020025>
- Romanuke, V. (2023). Random centroid initialization for improving centroid-based clustering. *Decision Making: Applications in Management and Engineering*, 6(2), 734–746. <https://doi.org/10.31181/dmame622023742>
- Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective Teacher–Student Relationships and Students’ Engagement and Achievement: A Meta-Analytic Update and Test of the Mediating Role of Engagement. *School Psychology Review*, 46(3), 239–261. <https://doi.org/10.17105/SPR-2017-0035.V46-3>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, 61, 101860. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Speybroeck, S., Kuppens, S., Van Damme, J., Van Petegem, P., Lamote, C., Boonen, T., & De Bilde, J. (2012). The Role of Teachers’ Expectations in the Association between Children’s SES and Performance in Kindergarten: A Moderated Mediation Analysis. *PLoS ONE*, 7(4), e34502. <https://doi.org/10.1371/journal.pone.0034502>
- Townsend, T. (2011). Thinking and acting both locally and globally: New issues for teacher education. *Journal of Education for Teaching*, 37(2), 121–137. <https://doi.org/10.1080/02607476.2011.558263>
- Wisniewski, B., Zierer, K., & Hattie, J. (2020). The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in Psychology*, 10, 3087. <https://doi.org/10.3389/fpsyg.2019.03087>
- Wolf, S., & McCoy, D. C. (2019). The role of executive function and social-emotional skills in the development of literacy and numeracy during preschool: A cross-lagged longitudinal study. *Developmental Science*, 22(4), e12800. <https://doi.org/10.1111/desc.12800>
- Yu, D., Zhou, X., Pan, Y., Niu, Z., & Sun, H. (2022). Application of Statistical K-Means Algorithm for University Academic Evaluation. *Entropy*, 24(7), 1004. <https://doi.org/10.3390/e24071004>
- Zhang, C., Yang, J., Li, M., & Deng, M. (2024). Simulation-Based Machine Learning for Predicting Academic Performance Using Big Data: *International Journal of Gaming and Computer-Mediated Simulations*, 16(1), 1–20. <https://doi.org/10.4018/IJGCMS.348052>
- Zinsser, K. M., & Curby, T. W. (2014). Understanding Preschool Teachers’ Emotional Support as a Function of Center Climate. *Sage Open*, 4(4), 2158244014560728. <https://doi.org/10.1177/2158244014560728>