Al-Ishlah: Jurnal Pendidikan
Vol.17, 3 (September, 2025), pp. 5119-5131
ISSN: 2087-9490 EISSN: 2597-940X, DOI: 10.35445/alishlah.v17i3.6210

Evaluating Mathematics Instruction under the Independent Curriculum: A Stake Model Analysis from Yogyakarta Junior High Schools

Nurhamidah Manullang[1], Jailani[2]

[1] Universitas Negeri Yogyakarta, Yogyakarta, Indonesia; nurhamidahmanullang.2021@student.uny.ac.id
[2] Universitas Negeri Yogyakarta, Yogyakarta, Indonesia; jailani@uny.ac.id

ARTICLE INFO

Keywords:

independent curriculum;
mathematics education;
curriculum evaluation;
Stake model;
differentiated instruction

Article history:

Received 2024-10-24
Revised 2025-02-18
Accepted 2025-09-30

ABSTRACT

The Independent Curriculum (Kurikulum Merdeka) in Indonesia aims to promote student-centered, differentiated learning. However, its implementation in mathematics education presents various challenges, particularly in planning, instruction, and assessment. This study employed a quantitative descriptive approach supported by qualitative data. The Stake evaluation model—comprising antecedents (planning), transactions (implementation), and outcomes (results)—was used to assess the curriculum's effectiveness in 12 junior high schools in Yogyakarta. Data were collected through document analysis of teaching modules and assessment tools, classroom observations, and evaluation of student learning outcomes. The planning phase showed that teaching modules in the independent sharing and changing categories were rated "good" (75.12% and 66.87%, respectively), while the independent learning category achieved a "very good" score (81.45%). In the implementation phase, independent sharing scored 75.90% ("good"), and both independent changing and learning categories scored 89.74% ("excellent"). However, in the assessment phase, scores were lower: 51.43% for independent sharing, 60.00% for independent changing, and 61.67% for independent learning, all falling into the "fairly good" category. While the planning and implementation stages align well with curriculum standards, assessment practices remain inconsistent and underdeveloped. The Independent Curriculum has been implemented adequately in mathematics learning across different categories, but the assessment component requires targeted improvement. Strengthening teacher capacity in assessment design and understanding of differentiated evaluation strategies is recommended to optimize learning outcomes.

Corresponding Author:
Nurhamidah Manullang
Universitas Negeri Yogyakarta, Yogyakarta, Indonesia; nurhamidahmanullang.2021@student.uny.ac.id

1. INTRODUCTION

Education plays a pivotal role in enhancing the quality of a nation's human resources (HR), which in turn influences its social and economic development. In Indonesia, however, the quality of education continues to face significant challenges. According to the World Population Review (2023), Indonesia

ranked 54th out of 78 countries in education in 2021, but this position declined to 67th out of 203 countries in 2022 and 2023. This downward trend underscores the urgent need for systemic reform to address both global competitiveness and local educational disparities.

In response to these challenges, the Indonesian Ministry of Education introduced the Independent Curriculum (Kurikulum Merdeka), which is designed to promote flexibility in learning, accommodate student diversity, and address learning disruptions caused by the COVID-19 pandemic (Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [Kemdikbud], 2022). The curriculum emphasizes contextual learning, the use of the Minimum Competency Assessment (AKM), and the cultivation of the Pancasila Student Profile—a framework aimed at fostering well-rounded, values-driven learners. Despite these progressive goals, the implementation of the Independent Curriculum—particularly in mathematics education—faces considerable obstacles in practice.

A continued decline in learning quality across several schools, highlighting the need for teachers to adopt more engaging and innovative teaching strategies (Education Report, 2023). One key recommendation is the integration of diverse learning media to support student understanding. Sholihatun (2021) found that the use of educational media significantly enhances student performance in mathematics. Nonetheless, field observations suggest that many educators are still in the adaptation phase. They report difficulties in executing differentiated learning, a central component of the Independent Curriculum. Teachers often perceive differentiation as time-consuming and challenging, especially in mathematics where managing varying skill levels requires strong classroom management and pedagogical competence.

Mathematics itself is a foundational subject that cultivates students' logical, critical, and analytical thinking (Munawaroh, 2017). Yet, various studies reveal that teachers encounter challenges in all phases of instructional practice—planning, implementation, and assessment. According to Lumbantoruan (2023), these challenges include: (1) inadequate lesson planning due to inconsistent preparation of the Learning Implementation Plan (RPP); (2) suboptimal delivery of instruction, especially with innovative strategies like Project-Based Learning (PjBL); and (3) insufficient alignment between assessment methods and the expected learning outcomes. Research also indicates that many teachers struggle to contextualize mathematics content in ways that are relevant to students' real-life experiences (Ismail, 2018; Kokotsaki, Menzies, & Wiggins, 2016).

Moreover, misconceptions about the value of instructional planning persist. Newsome (2017) found that some educators believe that lesson planning—including RPP development—has minimal impact on actual teaching performance. This perception often leads to poorly prepared instruction. In contrast, Fitriati (2023) argues that effective mathematics instruction requires systematic planning grounded in deep content knowledge and cognitive learning frameworks, such as Bloom's Taxonomy.

To comprehensively assess the effectiveness of the Independent Curriculum in mathematics education, evaluation must be conducted across three dimensions: planning, implementation, and assessment. The Stake Countenance Model, proposed by Stake (1967), provides a robust framework for such an evaluation. The model encompasses: (1) Antecedents—initial conditions and planning processes prior to implementation; (2) Transactions—the instructional activities and interactions during the learning process; and (3) Outcomes—the measurable results of the educational interventions.

By applying the Stake Model, this study aims to produce a comprehensive evaluation of the curriculum's implementation in mathematics learning. Specifically, it seeks to assess: (1) the quality of instructional planning, including the formulation of learning goals and lesson plans; (2) the effectiveness of instructional methods and technology use; and (3) the alignment between assessments and student learning needs. The findings are expected to inform evidence-based recommendations for improving the curriculum's implementation and ultimately contribute to enhancing the quality of education in Indonesia.

2. METHODS

This study employs a program evaluation research design utilizing a quantitative descriptive approach, complemented by qualitative data to support and enrich the findings. The primary focus of this evaluation is the implementation of mathematics learning in junior high schools within the city of Yogyakarta, specifically examining three core instructional components: planning, implementation, and assessment.

To guide the evaluation process, the Stake Countenance Model is adopted. This model, developed by Stake (1967) and later expanded by Fernandes (1984), emphasizes two critical components of evaluation: description and judgment. It categorizes the evaluation into three distinct stages: (1) Antecedents – conditions and planning activities prior to implementation, (2) Transactions – activities and interactions during the implementation process, and (3) Outcomes – the results achieved at the conclusion of the program. The Stake Model was selected as it aligns with the research objectives, which aim to systematically assess the quality of planning, the effectiveness of instructional delivery, and the appropriateness of assessments in mathematics learning under the Independent Curriculum (Kurikulum Merdeka). This model allows for a comprehensive and in-depth analysis, highlighting both the strengths and areas needing improvement at each stage of implementation.

Data collection in this study involved several instruments tailored to assess each evaluation component:

1. Observation Sheets: Classroom observations were conducted using structured observation sheets aligned with the three categories of curriculum implementation defined by the Ministry of Education—Independent Curriculum in the Independent Learning category, Independent Sharing category, and Independent Changing category. These sheets were used to record the real-time instructional practices of mathematics teachers.
2. Teaching Module Review: To evaluate the alignment between the teaching modules and the intended learning outcomes, a document analysis was performed using a dedicated module review sheet. This review ensured the coherence between curriculum content, learning objectives, and instructional strategies.
3. Assessment Document Review: The appropriateness and standard compliance of assessment tools were examined using an assessment evaluation sheet, which assessed whether the instruments were aligned with the competencies outlined in the Independent Curriculum.

To ensure the reliability of the instrument evaluations, particularly in rating observational and documentary data, interrater reliability was calculated. Three trained raters independently assessed the observation and document review sheets. The Fleiss' Kappa coefficient was used to measure the consistency among raters across the nine instruments used in this study, providing a robust estimate of agreement beyond chance (Gwet, 2014, p. 125).

The integration of both quantitative and qualitative data through structured observation, document review, and interrater reliability testing ensures that the evaluation results are valid, reliable, and comprehensive. This mixed-methods approach provides a nuanced understanding of how the Independent Curriculum is being implemented in mathematics classrooms and supports evidence-based recommendations for its improvement.

Table 1. Reliability Criteria for Research Instruments

No	Kappa Statistic	Explanation
1	< 0.40	Poor
2	0.40 – 0.75	Intermediate Good
3	> 0.75	Excellent

The following are the results of the reliability calculations for the nine instruments in the independent sharing category, the independent changing category, and the independent learning category:

Table 2. Interrater Reliability Coefficients of Evaluation Instruments Across Independent Curriculum Categories

Instrument	Reliability Coefficient (Fleiss' Kappa)	Interpretation
Independent sharing category teaching module document review sheet	0.74	Intermediate Good
Independent sharing category teaching module implementation sheet	0.94	Excellent
Independent sharing category assessment document review sheet	0.86	Excellent
Independent changing category teaching module document review sheet	0.89	Excellent
Independent changing category teaching module implementation sheet	0.81	Excellent
Independent changing category assessment document review sheet	0.88	Excellent
Independent learning category teaching module document review sheet	0.85	Excellent
Independent learning category teaching module implementation sheet	0.82	Excellent
Independent learning category assessment document review sheet	0.76	Excellent

Quantitative data is the main data used in this study. The data analysis technique used in this study is quantitative descriptive. Information collection is obtained through observation and documentation, which is analyzed using quantitative analysis. The quantitative data is analyzed using three stages, namely data scoring, data application, and data tabulation. This study's data analysis is based on the ideal mean and standard deviation. The results of the respondents' scores from each instrument are then used to interpret the material. The ideal mean is interpreted as a differentiator to determine the tendency of achievement scores to be on specific criteria. The criteria for processing quantitative data refer to the criteria put forward by (Azwar, 2010, pp.148).

Table 3. Quantitative Data Processing Criteria

No	Interval	Criteria
1	$M + 1.5\ SD < X \leq M + 3\ SD$	Excellent
2	$M + 0.5\ SD < X \leq M + 1.5\ SD$	Very Good
3	$M - 0.5\ SD < X \leq M + 0.5\ SD$	Good
4	$M - 1.5\ SD < X \leq M - 0.5\ SD$	Satisfactory
5	$M - 3\ SD < X \leq M - 1.5\ SD$	Poor

In this study, the antecedent aspect is measured through the components contained in the teaching module. The data collection technique used is a review of the teaching module documents for the independent sharing, changing, and learning categories. In the teaching module component for the independent sharing category, there are 86 statement items with a score range of 1 to 5. So, the highest ideal score is 430, and the lowest is 86. Thus, the ideal mean (M) = ½ (430+86) = 258 and the ideal standard deviation (SD) = 1/6 (430-86) = 57.33. The following is a table of evaluation criteria according to the ideal mean formula.

Table 4. Evaluation Criteria for Learning Planning in the Independent Sharing Category

No	Score X	Criteria
1	343.99 < X ≤ 429.99	Excellent
2	286.66 < X ≤ 343.99	Very Good
3	229.33 < X ≤ 286.66	Good
4	172.00 < X ≤ 229.33	Satisfactory
5	86.01 < X ≤ 172.00	Poor

The independent category teaching module component has 67 statement items with a score range of 1 to 5. So, the highest ideal score is 335, and the lowest is 67. Thus, the ideal mean (M) = ½ (335+67) = 201 and the ideal standard deviation (SD) = 1/6 (335-67) = 44.66. The following is a table of evaluation criteria according to the ideal mean formula.

Table 5. Evaluation Criteria for Independent Category Learning Planning Changes

No	Score X	Criteria
1	267.99 < X ≤ 334.98	Excellent
2	223.33 < X ≤ 267.99	Very Good
3	178.67 < X ≤ 223.33	Good
4	134.01 < X ≤ 178.67	Satisfactory
5	67.02 < X ≤ 134.01	Poor

In the independent learning category teaching module component, there are 55 statement items with a score range of 1 to 5. So, the highest ideal score is 275, and the lowest is 55. Thus, the ideal mean (M) = ½ (275+55) = 165 and the ideal standard deviation (SD) = 1/6 (275-55) = 36.67. The following is a table of evaluation criteria according to the ideal mean formula

Table 6. Evaluation Criteria for Independent Learning Category Learning Planning

No	Score X	Criteria
1	220.00 < X ≤ 275.01	Excellent
2	183.33 < X ≤ 220.00	Very Good
3	146.66 < X ≤ 183.33	Good
4	109.99 < X ≤ 146.66	Satisfactory
5	54.99 < X ≤ 109.99	Poor

In this study, the transaction aspect is measured through two components, namely, the implementation of learning process activities and the implementation of assessments. The technique for collecting data on learning process activities in the classroom is through observation. Observation of the learning process is based on the independent sharing category, the independent changing category, and the independent learning category. The implementation component of the independent sharing category has 39 statement items with a score range of 1 to 5. So the highest ideal score is 195 and the lowest ideal score is 39. Thus, the ideal mean (M) = ½ (195+39) = 117 and the ideal standard deviation (SD) = 1/6 (195-39) = 26. The following is a table of evaluation criteria according to the ideal mean formula.

Table 7. Evaluation Criteria for the Implementation of the Independent Sharing Learning Process Category

No	Score X	Criteria
1	156 < X ≤ 195	Excellent
2	130 < X ≤ 156	Very Good
3	104 < X ≤ 130	Good
4	78 < X ≤ 104	Satisfactory
5	39 < X ≤ 78	Poor

In the implementation component of the independent category changes, there are 39 statement items with a score range of 1 to 5. So, the highest ideal score is 195, and the lowest ideal score is 39. Thus, the ideal mean (M) = ½ (195+39) = 117 and the ideal standard deviation (SD) = 1/6 (195-39) = 26. The following is a table of evaluation criteria according to the ideal mean formula.

Table 8. Evaluation Criteria for the Implementation of the Independent Changing Learning Process Category

No	Score X	Criteria
1	156 < X ≤ 195	Excellent
2	130 < X ≤ 156	Very Good
3	104 < X ≤ 130	Good
4	78 < X ≤ 104	Satisfactory
5	39 < X ≤ 78	Poor

In the implementation component of the independent learning category, there are 31 statement items with a score range of 1 to 5. So, the highest ideal score is 155, and the lowest is 31. Thus, the ideal mean (M) = ½ (155+31) = 93 and the ideal standard deviation (SD) = 1/6 (155-31) = 20.67. The following is a table of evaluation criteria according to the ideal mean formula.

Table 9. Evaluation Criteria for the Implementation of the Learning Process in the Independent Learning Category

No	Score X	Criteria
1	124.00 < X ≤ 155.01	Excellent
2	103.33 < X ≤ 124.00	Very Good
3	82.66 < X ≤ 103.33	Good
4	61.99 < X ≤ 82.66	Satisfactory
5	30.99 < X ≤ 61.99	Poor

Furthermore, implementing learning assessments carried out by teachers in the classroom uses assessment documents. Data collection techniques use assessment document reviews. Document reviews are based on independent sharing categories, independent changing categories, and independent learning categories. In the implementation component of the independent sharing category assessment, there are statement items with a score range of 1 to 5. So, the highest ideal score is 35, and the lowest is 7. Thus, the ideal mean (M) = ½ (35+7) = 21 and the ideal standard deviation (SD) = 1/6 (35-7) = 4.67. The following is a table of evaluation criteria according to the ideal mean formula.

Table 10. Evaluation Criteria for the Implementation of Learning Assessments in the Independent Sharing Category

No	Score X	Criteria
1	28.00 < X ≤ 35.01	Excellent
2	23.33 < X ≤ 28.00	Very Good
3	18.66 < X ≤ 23.33	Good
4	13.99 < X ≤ 18.66	Satisfactory
5	6.99 < X ≤ 13.99	Poor

In the implementation component of the independent category assessment, there are statement items with a score range of 1 to 5. So, the highest ideal score is 35, and the lowest is 7. Thus, the ideal mean (M) = ½ (35+7) = 21 and the ideal standard deviation (SD) = 1/6 (35-7) = 4.67. The following is a table of evaluation criteria according to the ideal mean formula.

Table 11. Evaluation Criteria for the Implementation of Independent Category Learning Assessment Changes

No	Score X	Criteria
1	28.00 < X ≤ 35.01	Excellent
2	23.33 < X ≤ 28.00	Very Good
3	18.66 < X ≤ 23.33	Good
4	13.99 < X ≤ 18.66	Satisfactory
5	6.99 < X ≤ 13.99	Poor

In the implementation component of the independent learning category assessment, there are statement items with a score range of 1 to 5. So, the highest ideal score is 60, and the lowest is 12. Thus, the ideal mean (M) = ½ (60+12) = 36 and the ideal standard deviation (SD) = 1/6 (60-12) = 8. The following is a table of evaluation criteria according to the ideal mean formula.

Table 12. Evaluation Criteria for the Implementation of the Independent Learning Category Learning Assessment

No	Score X	Criteria
1	48 < X ≤ 60	Excellent
2	40 < X ≤ 48	Very Good
3	32 < X ≤ 40	Good
4	24 < X ≤ 32	Satisfactory
5	12 < X ≤ 24	Poor

In this study, the outcome aspect is measured through a review of the mid-semester assessment document of students. Then, the graduate competency standards of each school are compared with the scores obtained by teachers whose results are categorized as passing and failing from each school. The criteria for conducting the evaluation are used to analyze and interpret the results of the data processing obtained. The antecedent and transaction aspect criteria aim to analyze and interpret the results of data processing obtained in the field. The evaluation criteria refer to the ideal average criteria except for the outcome aspect. The outcome aspect criteria are converted using the following formula.

$$Percentage\ of\ achievement = \frac{average\ acquisition\ score}{the\ maximum\ ideal\ average\ of\ each\ aspect} \times 100\%$$

The average percentage of aspect scores to prove how much the implementation of aspects of junior high school mathematics learning in Yogyakarta City can be achieved. The percentage of the

total score is then used to interpret the suitability between the standards and the conditions obtained in the field. The percentage score of each aspect achievement is then compared with the success criteria in the following table.

Table 13. Success Criteria for the Independent Sharing Category Learning Program

Stage	Aspects	Success Criteria
Antecedent	Learning planning	100%
Transaction	Learning implementation	100%
	Assessment implementation	
Outcome	Student ASTS scores	83%

Table 14. Success Criteria for Independent Learning Program Category Change

Stage	Aspects	Success Criteria
Antecedent	Learning planning	100%
Transaction	Learning implementation	100%
	Assessment implementation	
Outcome	Student ASTS scores	84%

Table 15. Success Criteria for Independent Learning Program Category Change

Stage	Aspects	Success Criteria
Antecedent	Learning planning	100%
Transaction	Learning implementation	100%
	Assessment implementation	
Outcome	Student ASTS scores	83%

3. FINDINGS AND DISCUSSION

3.1 Research Findings

3.1.1 Antecedents (Planning Stage)

The quality of teaching modules prepared by teachers reflects the extent to which lesson planning aligns with the Independent Curriculum. Based on document analysis, teaching modules across the Independent Sharing, Independent Changing, and Independent Learning categories fall into the *good*, *quite good*, and *very good* criteria, with average percentages of 75.12%, 66.87%, and 81.45%, respectively. These findings indicate that, overall, teachers are capable of preparing modules that are largely aligned with the Process Standards outlined in Permendikbudristek No. 16 of 2022, which emphasize clarity, flexibility, and simplicity in planning.

However, some deficiencies were noted. In the Independent Sharing category, teachers often overlooked critical indicators such as learning outcomes, learning strategies, diagnostic assessments, and supporting materials (LKPD). Similarly, modules in the Independent Changing and Independent Learning categories lacked clarity in defining learning activities, learning resources, and appropriate assessment methods.

3.1.2 Transactions – Learning Implementation

Classroom observations were conducted to assess how well mathematics lessons were implemented according to the Independent Curriculum's differentiated instruction model.

Differentiation, as implemented, involved adjusting content, process, and products based on students' readiness and learning profiles.

The observation results show that classroom instruction in the Independent Sharing category was rated as *good* with a score of 75.90%, while both the Independent Changing and Independent Learning categories achieved *very good* ratings with identical scores of 89.74%. These findings suggest that most teachers have made efforts to apply differentiated instruction strategies. Nevertheless, issues such as the lack of student-centered learning activities, underutilization of constructivist approaches, and over-reliance on traditional teaching methods were still observed, especially in the Independent Sharing category.

3.1.3 Transactions – Assessment Implementation

The evaluation of assessment practices focused on the quality and compliance of assessment documents prepared by teachers. These were reviewed using standards set in Permendikbudristek No. 21 of 2022 and Permendikbud No. 66 of 2013, which emphasize diagnostic, formative, and summative assessments aligned with competencies.

Results show that the assessment implementation in the Independent Sharing, Independent Changing, and Independent Learning categories were rated as *not good*, *quite good*, and *quite good* respectively, with percentage scores of 51.43%, 60%, and 61.67%. Teachers frequently omitted diagnostic assessments, misaligned learning indicators with intended outcomes, and provided minimal feedback to students. These shortcomings significantly impact the quality of instruction and learner progression.

3.1.4 Outcomes (Student Achievement)

Student achievement was measured using the ASTS (Asesmen Sumatif Tengah Semester) scores and compared against the Minimum Completeness Criteria (KKTP). The findings show that in the Independent Sharing category, 81% of students did not meet the KKTP standard, while only 19% passed. In the Independent Changing category, 54.80% of students failed, and 45.20% passed. In the Independent Learning category, 66.50% failed and 33.50% passed.

These results indicate a significant gap between intended learning outcomes and actual student performance across all curriculum implementation categories.

3.1.5 Vertical and Horizontal Gaps Based on Stake's Model

Using Stake's Countenance Evaluation Model, a vertical comparison was conducted between the antecedent, transaction, and outcome stages. The ideal condition across all three stages was defined as 100% compliance. However, the actual observed data show substantial gaps, especially between the planning and assessment stages, and between implementation and student outcomes. These discrepancies suggest a lack of coherence across planning, instruction, and assessment, which may contribute to poor student performance.

3.2 Discussion

3.2.1 Lesson Planning Quality and Challenges

The results confirm that most teachers have made considerable efforts to prepare teaching modules that adhere to government standards. However, the inconsistencies in key indicators—such as learning objectives, strategies, and assessments—are concerning. This supports Küçükahmet's (2002) assertion that systematic planning is critical for effective teaching, helping to achieve instructional goals efficiently and with minimal resource expenditure. In contrast, Newsome (2017) found that some teachers perceive planning as less influential on actual teaching outcomes, which may explain the uneven quality observed in teaching modules.

A major challenge highlighted by this study is teachers' difficulty in operationalizing indicators and aligning them with measurable learning outcomes. This suggests the need for targeted professional

development on instructional design, particularly in constructing competency-based learning modules aligned with the Independent Curriculum.

3.2.2 Implementation of Differentiated Instruction

The high scores in classroom implementation for the Independent Changing and Independent Learning categories indicate that many teachers have embraced differentiated instruction principles. Differentiation in content, process, and product—as advocated by Tomlinson (2014)—can indeed help address diverse student needs. However, this study found that implementation often lacked constructivist learning elements, such as student agency and collaborative problem-solving, which are essential for deep learning in mathematics.

Sun (2019) notes that rigid instructional strategies discourage student engagement and participation, making learning less meaningful. In contrast, Cross (2009) argues that varied instructional methods can enhance student motivation and foster active learning. This aligns with findings from this study, which suggest that instructional monotony remains a challenge, particularly in the Independent Sharing category.

Further, many teachers tend to prioritize content coverage over process quality, resorting to conventional methods to meet curriculum demands. This behavior may stem from external pressures, such as standardized testing and time constraints, but it risks undermining the core principles of differentiated and student-centered learning.

3.2.3 Assessment Practices and Alignment Issues

Assessment implementation remains one of the most problematic aspects of curriculum delivery. The low scores across all categories, especially in the Independent Sharing group, reflect teachers' limited understanding of assessment types and their purposes. Research by Ulumudin and Fujianita (2019) and Wangid (2017) similarly found that many teachers struggle to conduct formative and diagnostic assessments effectively due to a lack of conceptual clarity and training.

Diagnostic assessments are critical for identifying students' prior knowledge and learning readiness, allowing for tailored instruction. However, this study revealed that such assessments are rarely conducted. Moreover, feedback—a key component of formative assessment—is often absent or ineffective, depriving students of actionable insights into their performance.

Improving assessment literacy among teachers is therefore imperative. This includes training in designing valid instruments, using rubrics, and interpreting assessment data to inform instruction.

3.2.4 Student Learning Outcomes and Influencing Factors

The ASTS results show that a significant number of students fail to meet the KKTP standards, particularly in the Independent Sharing category. This finding supports earlier research by Lukum (2015), who found that inadequate instructional quality contributes to poor student achievement. Moreover, while Wang, Han, Liu, and Xu (2021) emphasize the positive effects of innovative teaching strategies on learning outcomes, Sun (2019) found no direct correlation, suggesting that innovation alone is insufficient without proper alignment with curriculum standards.

Several contextual factors likely contribute to the low student performance observed in this study:

1. Low learning readiness: Many students lack the prerequisite knowledge and skills needed to grasp new mathematical concepts.
2. Student motivation: Mathematics is often perceived as a difficult subject, which lowers engagement and effort.
3. Teacher competency gaps: Many teachers do not fully understand how to implement the Independent Curriculum, especially regarding assessment and differentiated instruction.
4. Parental support: Limited communication and collaboration with parents further hinders student learning at home.

3.2.5 Structural and Policy-Level Challenges

In addition to pedagogical factors, structural challenges also impact curriculum implementation. The study found that many schools lack adequate facilities, such as reliable internet access and interactive learning technologies. These limitations hinder the integration of digital learning tools and project-based approaches, which are central to the Independent Curriculum's philosophy.

Furthermore, some schools implement the curriculum under pressure, without sufficient preparation or support from educational authorities. Teachers are often expected to adopt project-based learning (PjBL) and student-centered strategies with limited resources and professional development opportunities. This policy-driven implementation, without sufficient infrastructure and capacity building, risks superficial compliance rather than meaningful transformation.

3.2.6 Bridging the Gap Between Ideal and Actual Conditions

Stake's Countenance Evaluation Model provides a useful lens for analyzing the misalignment between planning (antecedent), instruction (transaction), and results (outcome). Ideally, these components should be mutually reinforcing, ensuring that well-planned lessons are effectively delivered and assessed, leading to improved student learning. However, this study reveals gaps at every stage:

1. Planning is generally good but lacks consistency in crucial indicators.
2. Implementation is improving but remains teacher-centered in many cases.
3. Assessment is poorly understood and poorly implemented.
4. Outcomes remain below expectations due to misalignment across stages.

Addressing these gaps requires a systemic approach, involving not only teacher training but also curricular reform, infrastructure development, and stronger community-school partnerships.

This study evaluated the implementation of mathematics learning under the Independent Curriculum using Stake's Countenance Model. Findings across antecedent, transaction, and outcome stages highlight significant strengths in teaching module preparation and classroom implementation, particularly in the Independent Changing and Independent Learning categories. However, assessment practices remain a weak point, and student achievement levels continue to fall below KKTP standards.

To improve implementation outcomes, several strategies are recommended:

1. Strengthen professional development programs focused on planning, differentiated instruction, and assessment literacy.
2. Enhance school infrastructure to support curriculum demands.
3. Foster collaboration among teachers, parents, and policy-makers to create a more supportive learning environment.
4. Reevaluate and gradually implement policies to ensure equitable curriculum adoption across all school contexts.

Ultimately, aligning planning, instruction, and assessment with student needs and curriculum goals will be critical for enhancing mathematics learning outcomes and overall educational quality in Indonesia.

4. CONCLUSION

Based on the results of the study, it can be concluded that in the evaluation of the planning variables of the antecedent stage of the independent sharing category, the independent changing category, and the independent learning category, the percentages of 75.12%, 66.87%, and 81.45% falls into the good, pleasing and very good criteria. Furthermore, the implementation variables of the transaction stage of the independent sharing category, the independent changing category, and the independent learning category, the percentages of 75.90%, 89.74%, and 89.74%, are in the good, excellent, and perfect criteria. Finally, the assessment variables of the transaction stage of the independent sharing category, the independent changing category, and the independent learning

category, the percentages of 51.43%, 60%, and 61.67%, are in the less good, quite good, and quite good criteria. In addition, the evaluation results at the outcome stage of the independent curriculum in the independent sharing category, the independent changing category, and the independent learning category, based on the ASTS value document, show that students tend to fail to meet the KKTP that has been set.

This study suggests that teachers develop and improve their own knowledge of differentiated learning, and assessment in the independent curriculum through teacher learning communities. The principal motivates teachers to develop better differentiated learning, especially related to assessment activities. In addition, this study is a reference for other subjects with a longer research duration so that the results can be more accurate. Then the government holds regular face-to-face training for teachers to monitor the development of teachers' understanding of the independent curriculum and conducts assessments on a national scale so that students are always motivated to learn.

REFERENCES

Cross, D. I. (2009). Alignment , Cohesion , and Change : Examining and Their Influence On Instructional Practices. *J Math Teacher Educ*, 325–346. https://doi.org/10.1007/s10857-009-9120-5

Fitriati, F. (2023). *Enhancing Prospective Mathematics Teachers ' Lesson Planning Skills Through Lesson Study Within School University Partnership Program*. *14*(1), 69–84.

Ismail, R. (2018). Perbandingan Keefektifan Pembelajaran Berbasis Proyek dan Pembelajaran Berbasis Masalah Ditinjau Dari Ketercapaian Tujuan Pembelajaran. *PYTHAGORAS: Jurnal Pendidikan Matematika*, *13*(2), 181–188. https://doi.org/10.21831/pg.v13i2.23595

Kemdikbud. (2022). Buku Saku Kurikulum Merdeka; Tanya Jawab. *Kementerian Pendidikan Dan Kebudayaan*, 1–50.

Kokotsaki. (2016). Project-based learning: A review of the literature. *Improving Schools*, 1–11. https://doi.org/10.1177/1365480216659733

Lee, H. Y., Lin, C. J., Wang, W. S., Chang, W. C., & Huang, Y. M. (2023). Precision education via timely intervention in K - 12 computer programming course to enhance programming skill and affective - domain learning objectives. *International Journal of STEM Education*. https://doi.org/10.1186/s40594-023-00444-5

Lukum, A. (2015). Evaluasi Program Pembelajaran Ipa Smp Menggunakan Model Countenance Stake. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *19*(1), 25–37. https://doi.org/10.21831/pep.v19i1.4552

Lumbantoruan. (2023). Kesulitan Guru Dalam Mengimplementasikan Kurikulum Merdeka Belajar Materi Matematika di Sekolah Menegah Pertama (SMP). *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *12*(1), 1459–1473. https://doi.org/10.24127/ajpm.v12i1.7082

Munawaroh. (2017). The Influence Of Teaching Methods and Learning Environment to The Student's Learning Achievement Of Craft and Entrepreneurship Subjects at Vocational High School. *International Journal of Environmental & Science Education*, *12*(4), 665–678. Retrieved from https://files.eric.ed.gov/fulltext/EJ1144859.pdf

Newsome. (2017). Teacher Pedagogical Content Knowledge , Practice , and Student Achievement. *International Journal of Science Education*, 1–20. https://doi.org/10.1080/09500693.2016.1265158

Sholihatun. (2021). Pengembangan Media Pembelajaran Bangun Ruang Sisi Datar Menggunakan Macromedia Flash 8 Berbasis Pendekatan PMRI. *PYTHAGORAS: Jurnal Pendidikan Matematika*, *16*(2), 189–203. https://doi.org/10.21831/pythagoras.v16i2.42194

Sibaen. (2023). Principal Components Of Students ' Difficulties In Mathematics In The Purview Of Flexible Learning. *Journal on Mathematics Education*, *14*(2), 353–374.

Stake, R. E. (1967). The Countenance of Educational Evaluation. *Teachers College Record: The Voice of Scholarship in Education*, *68*(7), 1–15. https://doi.org/10.1177/016146816706800707

Sun, K. L. (2019). The mindset disconnect in mathematics teaching: A qualitative analysis of classroom instruction. *Journal of Mathematical Behavior*, *56*(April), 1. https://doi.org/10.1016/j.jmathb.2019.04.005

Ulumudin, I., & Fujianita, S. (2019). The implementation of attitude assessment in Curriculum 2013 at elementary schools. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *23*(1), 46–56. https://doi.org/10.21831/pep.v23i1.23391

Wang, R., Han, J., Liu, C., & Xu, H. (2021). How Do University Students' Perceptions of the Instructor's Role Influence Their Learning Outcomes and Satisfaction in Cloud-Based Virtual Classrooms During the COVID-19 Pandemic? *Frontiers in Psychology*, *12*(April), 1–12. https://doi.org/10.3389/fpsyg.2021.627443

Wangid. (2017). The evaluation of authentic assessment implementation of Curriculum 2013 in Elementary School. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *21*(1), 104–115. https://doi.org/10.21831/pep.v21i1.15779