

Vespera Studia Critica (VSC)
Vol 1 No 1 October 2026, Hal 27-40
ISSN: XXXX-XXXX (Print) ISSN: XXXX-XXXX (Electronic)
Open Access: https://scriptaintelektual.com/scripta-humanika/index

Indonesia's Learning Decline: A Longitudinal Reappraisal of PISA Performance and Minimum Proficiency

Ayu Ambarwati[1*]
[1] Politeknik Bina Madani, Indonesia
email: ambarwatiayuu1999@gmail.com[1]

Article Info :
Received:
05-09-2026
Revised:
12-10-2026
Accepted:
21-10-2026

Abstract

Indonesia's persistent learning challenges raise concerns about the extent to which educational participation translates into measurable foundational competencies. This study examined longitudinal changes in mathematics, reading, and science achievement across the 2012, 2015, 2018, and 2022 Programme for International Student Assessment (PISA) cycles and evaluated minimum-proficiency attainment in 2022. A quantitative descriptive longitudinal design was employed using 12 aggregate observations obtained from official OECD publications. Data were analyzed through consecutive-cycle comparisons, endpoint score differences, and descriptive calculations of proficiency distributions. Reading performance declined from 396 to 359 points, representing a 37-point reduction, while mathematics decreased from 375 to 366 points. Science increased marginally from 382 to 383 points across the full period, despite declining after its 2015 peak. Reading deterioration was evident before the COVID-19 pandemic. In 2022, only 18% of Indonesian students achieved at least Level 2 proficiency in mathematics, 25% in reading, and 34% in science. These findings demonstrate domain-specific achievement trajectories and substantial foundational proficiency deficits. The study contributes an integrated descriptive assessment supporting differentiated educational remediation and evidence-based evaluation, while acknowledging that aggregate indicators cannot establish causal relationships or policy effectiveness.

Keywords: *Educational equity; Indonesia; Learning outcomes; Minimum proficiency; PISA.*

©2026 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.
(https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUCTION

Educational participation and measurable learning achievement have become increasingly disconnected policy concerns, particularly in education systems where expanded schooling has not consistently produced improvements in foundational competencies. The Programme for International Student Assessment (PISA) provides a comparative framework for examining this discrepancy through students' mathematical, reading, and scientific literacy, although its aggregate indicators cannot independently establish the causes of educational performance. Indonesia represents a significant case because its educational development has been accompanied by persistent difficulties in achieving internationally defined proficiency standards. The PISA 2022 assessment reported that only 18% of Indonesian students attained at least Level 2 proficiency in mathematics, 25% in reading, and 34% in science, indicating substantial limitations in the competencies required to interpret information, solve familiar problems, and apply scientific reasoning (Organisation for Economic Co-operation and Development [OECD], 2023). Recent Indonesian research has increasingly examined the relationship between educational achievement, socioeconomic disadvantage, school governance, and institutional learning conditions, suggesting that performance disparities cannot be adequately understood through national averages alone (Aditia & Széll, 2025; Miftakhi & Waluyo, 2025). Evidence from Southeast Asian reading assessments further demonstrates the relevance of student- and school-level characteristics to the interpretation of educational achievement (Darmawan & Dharmapatni, 2024). Comparative research addressing educational policy and learning culture has also emphasized differences in national educational environments that may influence the interpretation of PISA outcomes (Purnasih et al., 2026). Concerns surrounding learning deterioration have intensified following pandemic-related educational disruptions, with Indonesian studies examining learning loss during school closures and subsequent transitions in instructional

arrangements (Fatayan et al., 2023; Wibowo, 2022). These developments establish the need to distinguish persistent educational weaknesses from more recent disruptions while examining whether changes in national achievement correspond to improvements in minimum learning proficiency.

Contemporary scholarship on international educational assessment has developed through several interconnected research directions involving measurement validity, achievement determinants, socioeconomic inequality, and instructional conditions. Studies using PISA data have examined the association between information and communication technology competencies and performance in mathematics, science, and reading, while investigations of mathematical learning strategies have demonstrated the methodological value of examining learning-related characteristics rather than treating achievement scores as isolated indicators (Acun Çelik et al., 2024; Kim et al., 2024). Research on reading assessment has similarly extended beyond mean performance comparisons to consider item difficulty, metacognitive strategies, and the relationship between educational environments and literacy development (Ghimire & Mokhtari, 2025; Marcq & Braeken, 2025). A parallel body of scholarship has challenged narrowly defined educational equity measures by advocating multidimensional interpretations of achievement differences and examining inequalities within national educational systems (Appels et al., 2024; Cito & Marôco, 2026). Investigations using PISA 2018 and 2022 have connected scientific literacy with socioeconomic differences, while cross-national analyses have explored relationships among scientific competencies, attitudes, values, and global competence (Kastorff & Heine, 2025; Kastorff et al., 2026). Research addressing socioeconomic inequalities in reading literacy has further expanded the discussion from performance differences toward learning-related characteristics associated with those differences (Yang, 2026). Methodological studies have introduced another essential consideration by investigating unequal patterns of insufficient-effort responding across student groups, schools, and cultures, thereby demonstrating that the interpretation of large-scale assessment outcomes depends partly on measurement quality (Wang, 2025). Collectively, these developments indicate a transition from descriptive international rankings toward more differentiated interpretations of achievement, equity, measurement, and learning environments. Their collective contribution establishes a strong analytical foundation for examining Indonesia's educational performance, although the specific relationship between long-term national achievement trajectories and minimum proficiency requires a distinct descriptive assessment.

Despite substantial progress in educational assessment research, several limitations remain relevant to the interpretation of Indonesia's learning trajectory. Research on educational inequality has demonstrated the importance of institutional differences and cross-national socioeconomic conditions, but comparisons of school effectiveness and national educational disparities address analytical questions that differ from reconstructing changes in a single country's achievement across successive PISA cycles (Steinmann & Olsen, 2022; Maheshwari & Koria, 2025). Investigations of instructional quality and socioeconomic disparities provide more differentiated explanations of achievement variation, yet their explanatory designs depend on information that cannot be recovered from national mean scores alone (Patsis et al., 2026). Research examining digital access, digital use, and mathematical achievement similarly highlights potentially relevant educational conditions without establishing how the relative performance of mathematics, reading, and science evolved across Indonesia's 2012–2022 assessment sequence (Özdoğan Özbal, 2026). Studies employing propensity score matching to evaluate grade retention demonstrate the methodological demands of estimating educational effects, reinforcing the distinction between explanatory research and descriptive comparisons based on published aggregate indicators (León & Martínez-Abad, 2025). Pandemic-related scholarship introduces an additional interpretive challenge. A systematic review and meta-analysis documented learning disruption during COVID-19, while studies of differential achievement changes, socio-emotional skills, and teacher support examined dimensions of educational deterioration associated with the pandemic period (Betthäuser et al., 2023; Castellano et al., 2025; Lichand et al., 2024; Golle et al., 2026). These contributions are valuable but cannot independently determine whether Indonesia's observed weaknesses originated before the pandemic or followed comparable trajectories across academic domains. The research gap addressed here is consequently a bounded empirical and methodological problem: the need for an integrated, transparent account of Indonesia's four-cycle PISA performance that distinguishes domain-specific achievement changes,

pre-pandemic developments, and the distribution of students below minimum proficiency without converting descriptive associations into unsupported causal explanations.

Failure to address this interpretive gap has consequences for both educational scholarship and the formulation of evidence-based improvement strategies. National performance deterioration may be interpreted primarily as a consequence of recent educational disruptions when earlier changes in achievement remain insufficiently distinguished from subsequent developments. Such an interpretation could encourage policy responses that address temporary instructional interruption without adequately considering persistent deficiencies in foundational literacy, numeracy, and scientific reasoning. The relationship between achievement indicators and educational intervention is especially consequential because instructional strategies differ in their learning demands and implementation requirements. Research examining the balance between teacher-led and student-led science instruction emphasizes cognitive load as a relevant consideration in the design of learning activities (Cairns, 2025). Investigations of intensive supplemental reading and mathematics programs demonstrate the relevance of evaluating targeted instructional support over time rather than assuming that additional educational provision necessarily improves measurable competencies (Lindström-Sandahl et al., 2025). Reading research also identifies teacher feedback and students' reading self-efficacy as relevant dimensions of educational practice, while studies of metacognitive strategies provide a complementary basis for considering how reading development may be supported (Wang et al., 2025; Ghimire & Mokhtari, 2025). Indonesian evidence concerning teachers' digital literacy and student learning loss further illustrates the need to consider instructional capacity when discussing educational recovery (Safrizal et al., 2023). These findings do not establish that particular interventions caused changes in Indonesia's national PISA scores. They indicate that policy recommendations require a clear distinction between observed achievement patterns and mechanisms that must be evaluated independently. A systematic description of long-term performance and proficiency thresholds can support more appropriate prioritization of foundational learning, while preventing national assessment results from being treated as direct evidence of curriculum effectiveness, teacher performance, or individual learning trajectories.

The present study is positioned at the intersection of longitudinal educational assessment, foundational learning proficiency, and evidence-based interpretation of national education indicators. Its contribution differs from research primarily concerned with socioeconomic achievement determinants, school governance, or learning environments because it does not attempt to estimate relationships between individual characteristics and educational outcomes (Aditia & Széll, 2025; Miftakhi & Waluyo, 2025). It also differs from comparative analyses that examine educational inequality across countries or assessment systems, since its analytical unit is Indonesia's published national performance across four PISA cycles and three competency domains (Appels et al., 2024; Maheshwari & Koria, 2025; Purnasih et al., 2026). Whereas pandemic-related research provides evidence concerning learning disruption, this investigation places the 2018–2022 interval within the longer 2012–2022 assessment sequence to avoid interpreting recent score changes without reference to preceding developments (Betthäuser et al., 2023; Castellano et al., 2025). The proposed novelty is not the introduction of a new educational theory, the discovery of previously unavailable PISA observations, or the identification of causal determinants. Its more limited but identifiable contribution consists of integrating three descriptive components within a single reproducible analytical framework: domain-specific national score trajectories, absolute differences between consecutive assessment cycles, and the proportions of students reaching or remaining below minimum proficiency in the latest assessment. This integration establishes a distinction between average achievement changes and proficiency distributions, which convey related but analytically different information about educational performance. The approach also incorporates methodological caution concerning the interpretation of aggregated international assessment data, consistent with research emphasizing measurement characteristics and data quality in large-scale educational evaluations (Marcq & Braeken, 2025; Wang, 2025). The originality claimed is therefore a transparent and context-specific synthesis of existing indicators rather than a new statistical dataset or explanatory model. Its scientific value depends on the consistency of its calculations, the relevance of its comparisons, and the explicit recognition of the limits of descriptive inference.

This study aims to examine Indonesia's educational performance across the 2012, 2015, 2018, and 2022 PISA assessment cycles by comparing national mean scores in mathematics, reading, and

science and evaluating the proportions of students achieving minimum proficiency in the latest assessment. Specifically, the analysis seeks to identify the magnitude and direction of achievement changes between consecutive assessment periods, determine differences in long-term performance trajectories across the three competency domains, and describe the distribution of students above and below the established minimum proficiency threshold. The research employs a descriptive longitudinal approach using publicly available aggregate assessment statistics, with calculations limited to reported indicators and directly derived numerical differences. Its scientific contribution lies in providing an integrated interpretation of national achievement patterns that distinguishes changes in average performance from the distribution of foundational competencies while avoiding unsupported claims concerning educational causation. Its methodological contribution consists of establishing an explicit and reproducible comparison procedure that can be evaluated against the original aggregate indicators without requiring simulated observations or unreported individual-level data. Its practical contribution is to provide a clearer empirical basis for identifying priority areas in foundational education, differentiated learning support, and subsequent institutional evaluation. The study is intended to support a more precise interpretation of Indonesia's educational challenges and to identify areas where further student-level, school-level, and intervention-based research is required to explain achievement differences and evaluate potential improvements.

RESEARCH METHODS

This study employed a quantitative descriptive longitudinal design using secondary data obtained from the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA). The research examined changes in Indonesia's educational performance across four assessment cycles conducted in 2012, 2015, 2018, and 2022. The unit of analysis was the nationally aggregated mean achievement score for each assessment year and competency domain, comprising mathematics, reading, and science. The analytical dataset consisted of 12 observations representing three competency domains across four assessment cycles. Data were obtained exclusively from officially published OECD assessment reports, including the PISA 2018 results, PISA 2022 results, and the Indonesia country report accompanying the 2022 assessment. The selection of assessment cycles was determined by the study's predefined observation period and the availability of comparable aggregate indicators, rather than by probability sampling or the recruitment of individual participants. Data collection involved documentary extraction of published national achievement scores and minimum-proficiency percentages, with particular attention to preserving the original assessment years, competency classifications, measurement units, and indicator definitions. The analysis incorporated the percentage of Indonesian students achieving at least Level 2 proficiency in mathematics, reading, and science in 2022, together with the complementary proportions falling below that threshold. The study did not employ questionnaires, interviews, classroom observations, experimental interventions, or individual-level student records. Its longitudinal classification referred to repeated national assessments across successive years rather than repeated measurements of the same students.

Data analysis was conducted using descriptive comparisons of national achievement scores and minimum-proficiency distributions. Achievement changes were calculated separately for mathematics, reading, and science by subtracting the earlier assessment score from the subsequent assessment score, expressed as $\Delta S = S_t - S_{t-1}$, where S represents the published national mean score and t denotes the assessment cycle. This procedure was applied to the consecutive intervals 2012–2015, 2015–2018, and 2018–2022, as well as to the full observation period from 2012 to 2022. Positive differences indicated increases in estimated national mean achievement, whereas negative differences indicated decreases. Minimum-proficiency analysis used the percentages of students attaining Level 2 or higher in the 2022 assessment, while the proportion below Level 2 was calculated by subtracting the reported attainment percentage from 100%. Results were organized into comparative tables and graphical representations to distinguish achievement trajectories across competency domains and differences between average performance and proficiency distributions. Analytical accuracy was addressed by retaining the OECD's original indicator definitions, separating directly reported values from derived arithmetic calculations, and maintaining consistency between the numerical tables and their accompanying interpretations. Comparisons were restricted to descriptive changes in published aggregate estimates; differences were not interpreted as statistically

significant effects because standard errors and inferential testing were not incorporated into the analysis. The study did not estimate causal relationships, evaluate the effects of educational policies, or attribute observed achievement changes to the COVID-19 pandemic or other institutional factors. Differences in assessment populations across cycles were acknowledged as a limitation of aggregate longitudinal comparison. These methodological boundaries ensured that the interpretation remained consistent with the available secondary data and that the calculations could be independently reproduced using the specified OECD publications.

RESULTS AND DISCUSSION

Long-Term Performance Across Mathematics, Reading, and Science

Indonesia's educational performance between 2012 and 2022 exhibited distinct patterns across mathematics, reading, and science, indicating that changes in national learning achievement cannot be adequately represented by a single aggregate interpretation. The four PISA assessment cycles provide a longitudinal basis for examining differences in achievement trajectories, although the reported scores represent successive national samples rather than repeated observations of identical students. Mathematics and reading demonstrated negative endpoint changes, whereas science displayed a marginal positive difference despite substantial fluctuations during the observation period. These differences warrant domain-specific interpretation because international educational assessments measure competencies that involve different cognitive demands, learning strategies, and assessment characteristics (Acun Çelik et al., 2024; Kim et al., 2024).

Reading performance represents a particularly consequential dimension of Indonesia's educational trajectory because reading literacy supports the interpretation, evaluation, and application of information across academic disciplines. The national mean reading score was 396 in 2012 and increased marginally to 397 in 2015, before declining to 371 in 2018 and 359 in 2022. This sequence indicates that the deterioration in reading achievement had already emerged before the pandemic-related disruption of formal schooling, challenging interpretations that associate the observed decline exclusively with educational conditions after 2020. Research on reading assessment has emphasized that achievement outcomes reflect differences in the cognitive and measurement demands of reading tasks, making careful interpretation of proficiency indicators necessary when evaluating educational performance (Marcq & Braeken, 2025).

Mathematics demonstrated a different trajectory, characterized by an initial improvement followed by successive reductions in national achievement scores. Indonesia's mathematics mean increased from 375 in 2012 to 386 in 2015, before declining to 379 in 2018 and 366 in 2022, suggesting that the improvement recorded during the earlier assessment interval was not sustained. The observed changes are relevant to research examining mathematical learning strategies and the relationship between students' cognitive approaches and achievement, although the available aggregate data cannot identify the instructional factors responsible for the fluctuations (Kim et al., 2024). The decline from the 2015 peak indicates that endpoint comparisons alone provide an incomplete account of educational performance, particularly when intermediate assessment cycles reveal substantial differences in achievement direction.

Science achievement followed a less uniform pattern, with the national mean increasing from 382 in 2012 to 403 in 2015 before declining to 396 in 2018 and 383 in 2022. The one-point positive difference between the initial and final assessments contrasts with the substantial reduction from the highest observed score, demonstrating how the selection of comparison years can influence interpretations of educational progress. Science literacy involves the application of scientific knowledge, interpretation of evidence, and evaluation of scientific explanations, making its assessment conceptually distinguishable from reading and mathematical proficiency. Research examining scientific literacy and related competencies further indicates that achievement should be interpreted in relation to the specific knowledge and reasoning demands represented by assessment indicators rather than treated as a general measure of educational quality (Kastorff et al., 2026).

The comparative data presented in Table 1 establish the numerical basis for evaluating achievement changes across the three competency domains. Each observation represents an officially published national mean score corresponding to a particular PISA assessment year, without the introduction of simulated student records or estimated intermediate values. The table preserves the original chronological sequence to distinguish initial performance, intermediate fluctuations, and final

achievement levels. This presentation supports transparent comparison while maintaining the distinction between descriptive national statistics and explanatory analyses requiring individual- or school-level information (OECD, 2023; Darmawan & Dharmapatni, 2024).

Table 1. Indonesia's PISA Achievement Scores Across Assessment Cycles, 2012–2022

Assessment year	Mathematics	Reading	Science
2012	375	396	382
2015	386	397	403
2018	379	371	396
2022	366	359	383

Source: OECD PISA aggregate assessment statistics, 2012–2022.

The observations in Table 1 reveal that Indonesia's educational achievement did not decline at a constant rate across successive assessment periods. Reading exhibited a comparatively stable initial interval followed by substantial reductions, mathematics experienced an earlier improvement before reversing direction, and science displayed a pronounced increase followed by two consecutive declines. These differences demonstrate that educational performance deterioration must be examined through the temporal characteristics of individual competency domains rather than inferred from the direction of the final assessment alone. The analytical importance of preserving domain-specific patterns is consistent with research demonstrating that large-scale assessments capture different dimensions of learning and that student characteristics may be associated differently with achievement across subjects (Acun Çelik et al., 2024; Darmawan & Dharmapatni, 2024).

The magnitude of achievement changes further clarifies the differences between short-term fluctuations and longer-term educational trajectories. Between 2012 and 2015, mathematics increased by 11 points, reading by one point, and science by 21 points, whereas the subsequent 2015–2018 interval produced declines of seven, 26, and seven points, respectively. The final interval from 2018 to 2022 recorded reductions of 13 points in mathematics, 12 points in reading, and 13 points in science, demonstrating that negative changes occurred across all three domains during the latest observed assessment period. Research on educational disruption during COVID-19 provides relevant comparative evidence of learning deterioration, but its findings do not establish that the reductions observed in Indonesia were exclusively attributable to the pandemic (Betthäuser et al., 2023; Castellano et al., 2025).

The distinction between pre-pandemic deterioration and subsequent educational disruption is especially relevant to interpreting Indonesia's reading performance. The 26-point reduction between 2015 and 2018 demonstrates that substantial weakening in reading achievement preceded the school closures associated with COVID-19, whereas the additional reduction between 2018 and 2022 occurred across a period that included major disruptions to educational provision. Existing research has examined how school closures, instructional support, and changes in learning conditions were associated with achievement and motivation, although the magnitude and distribution of these relationships may differ across educational settings (Golle et al., 2026; Lichand et al., 2024). The national PISA series cannot disentangle these mechanisms because it does not incorporate the explanatory variables necessary to distinguish pandemic effects from pre-existing institutional and instructional conditions.

The interpretation of longitudinal achievement patterns also requires methodological caution regarding the comparability and substantive meaning of changes in published national estimates. Differences between PISA cycles describe changes in estimated achievement among successive nationally assessed student populations and should not be interpreted as individual learning gains or losses. Research on response behavior in large-scale assessments has emphasized the relevance of data-quality differences across students, schools, and cultural environments, reinforcing the need to avoid treating observed mean differences as direct measures of instructional effectiveness (Wang, 2025). The available dataset does not include standard errors, student-level characteristics, or school-level explanatory variables, preventing statistical significance testing and causal attribution within the present analytical design.

The longitudinal evidence identifies reading as the domain exhibiting the largest negative endpoint difference, with a 37-point reduction between 2012 and 2022, compared with a nine-point reduction in mathematics and a one-point increase in science. These differences indicate that the direction and magnitude of change depend on the competency examined and the assessment interval selected, rather than reflecting an identical pattern of deterioration across all academic domains. Research on socioeconomic differences in scientific literacy illustrates the importance of distinguishing aggregate achievement from the conditions underlying variation among students, although these relationships cannot be estimated using the present national mean series (Kastorff & Heine, 2025). The findings establish a descriptive basis for investigating Indonesia's minimum-proficiency distribution and educational inequality, particularly whether differences in average national performance are accompanied by substantial proportions of students who have not attained foundational competency thresholds.

Minimum Proficiency and Educational Inequality in Indonesia

Minimum proficiency provides an essential analytical dimension for evaluating educational achievement because national mean scores do not indicate how many students possess the competencies required to perform fundamental academic tasks. The PISA 2022 assessment identified substantial proportions of Indonesian students who had not reached Level 2 proficiency in mathematics, reading, and science, revealing educational challenges extending beyond changes in average performance. The distinction between mean achievement and proficiency distribution is particularly relevant because improvements in national averages do not necessarily imply equivalent progress among students performing below established competency thresholds. Research on educational equity has emphasized the limitations of relying exclusively on aggregate achievement indicators, supporting the consideration of distributional outcomes when evaluating differences in educational opportunity (Appels et al., 2024; Steinmann & Olsen, 2022).

Mathematical proficiency represents a critical concern because Level 2 performance reflects the capacity to apply basic mathematical reasoning in familiar situations and interpret relatively straightforward quantitative information. The OECD reported that 18% of Indonesian students attained at least Level 2 in mathematics in 2022, leaving 82% below the established proficiency threshold. This distribution indicates that a substantial majority of assessed students had not demonstrated the level of mathematical competency represented by the international baseline, although it does not establish that these students lacked every form of mathematical knowledge. Research examining socioeconomic inequality and mathematics achievement has highlighted the importance of distinguishing access to educational resources from the effective development of mathematical competencies (Özdoğan Özbal, 2026).

Reading proficiency presents a related challenge because the ability to understand written information supports learning across academic disciplines and influences students' capacity to engage with increasingly complex instructional materials. In 2022, 25% of Indonesian students achieved at least Level 2 in reading, while 75% remained below the threshold, indicating that minimum reading competency was not demonstrated by most assessed students. This finding is particularly consequential when considered alongside the decline in Indonesia's national mean reading score between 2012 and 2022, although the two indicators describe different characteristics of educational performance. Research examining socioeconomic inequalities in reading literacy has demonstrated the relevance of learning-related characteristics and educational disadvantage to the interpretation of literacy outcomes (Yang, 2026; Aditia & Széll, 2025).

Scientific literacy exhibited a comparatively higher minimum-proficiency attainment rate, although the proportion of students below Level 2 remained substantial. The PISA 2022 assessment reported that 34% of Indonesian students achieved the minimum science proficiency threshold, whereas 66% did not reach that level. The higher attainment percentage relative to mathematics and reading should not be interpreted as evidence that science education is institutionally more effective, because the three domains involve different competency frameworks and assessment characteristics. Research investigating scientific literacy and socioeconomic disparities emphasizes the relevance of educational conditions to differences in achievement, while comparative studies of scientific competencies demonstrate that scientific reasoning should be evaluated through its specific assessment constructs (Kastorff & Heine, 2025; Kastorff et al., 2026).

Table 2 presents the distribution of Indonesian students according to minimum proficiency attainment in the three PISA competency domains. The percentages attaining Level 2 or higher were obtained from the published PISA 2022 Indonesia indicators, while the corresponding percentages below Level 2 were calculated as complementary proportions. These values allow the extent of minimum-proficiency attainment to be compared across academic domains without introducing hypothetical student observations or additional statistical estimates. The distinction between attainment and non-attainment is consistent with research advocating multidimensional approaches to educational equity and cautioning against reducing educational performance to a single summary indicator (Appels et al., 2024; Cito & Marôco, 2026).

Table 2. Distribution of Indonesian Students by Minimum Proficiency in PISA 2022

Competency domain	At or above Level 2 (%)	Below Level 2 (%)
Mathematics	18	82
Reading	25	75
Science	34	66

Source: OECD (2023), PISA 2022 Indonesia country indicators.

The distribution presented in Table 2 indicates that the proportion of students below minimum proficiency exceeded the proportion meeting the threshold in every assessed domain. Mathematics exhibited the largest below-threshold percentage at 82%, followed by reading at 75% and science at 66%, producing descriptive differences of seven percentage points between mathematics and reading and nine percentage points between reading and science. These differences identify variation in the prevalence of minimum-proficiency attainment but do not establish that the underlying competency scales are directly interchangeable across subjects. Research on educational inequality has demonstrated that differences in achievement distributions require interpretation alongside broader educational conditions, rather than being treated as direct evidence of the effectiveness of particular schools or policies (Maheshwari & Koria, 2025; Miftakhi & Waluyo, 2025).

The relationship between minimum proficiency and educational inequality requires particular attention because students' opportunities to acquire foundational competencies may differ according to household circumstances, school characteristics, and access to instructional support. Previous research involving Indonesian educational achievement has investigated the relationship between school governance, contextual disadvantage, and performance differences, indicating that national results can conceal substantial variation in educational conditions (Aditia & Széll, 2025; Miftakhi & Waluyo, 2025). Comparative evidence from Southeast Asian reading assessments similarly supports the analytical relevance of student- and school-level characteristics when investigating literacy achievement (Darmawan & Dharmapatni, 2024). The present dataset does not contain the socioeconomic classifications or institutional covariates necessary to measure these relationships, meaning that the observed proficiency deficits establish the scale of the achievement problem rather than the distribution of its causes across disadvantaged groups.

Educational inequality also has an instructional dimension because differences in teaching quality and learning support may affect students' opportunities to develop the competencies represented by minimum-proficiency thresholds. Research investigating socioeconomic disparities in mathematics achievement has examined instructional quality as a potentially relevant dimension of educational inequality, while studies of mathematics learning strategies have demonstrated the importance of considering students' approaches to mathematical tasks (Patsis et al., 2026; Kim et al., 2024). These findings provide a conceptual basis for distinguishing low measured proficiency from the institutional mechanisms that may contribute to its persistence. The Indonesian PISA aggregate indicators cannot determine whether the below-threshold proportions resulted from instructional practices, learning resources, socioeconomic conditions, or interactions among these factors, and such explanations require additional evidence beyond the descriptive comparisons undertaken here.

The substantial proportion of students below minimum proficiency also raises questions about the relationship between educational disruption and the continuity of foundational learning. Indonesian studies have documented learning difficulties associated with pandemic-era educational arrangements, including learning from home and challenges involving teachers' digital literacy, while

international research has examined the broader consequences of disrupted schooling for educational achievement (Fatayan et al., 2023; Safrizal et al., 2023; Wibowo, 2022; Betthäuser et al., 2023). These studies establish the relevance of instructional continuity to educational recovery but do not permit the attribution of Indonesia's 2022 proficiency distribution exclusively to pandemic-related circumstances. The persistence of achievement concerns across earlier PISA cycles indicates the need to consider long-term educational performance when developing explanations for low proficiency attainment. Evidence from the present analysis supports prioritizing foundational learning as a policy concern while leaving the effectiveness of specific interventions and the relative contribution of educational disruptions to subsequent causal research.

The interpretation of minimum-proficiency outcomes must also recognize the measurement boundaries of international educational assessment. Level 2 represents an assessment-defined benchmark rather than a comprehensive classification of students' intellectual ability, literacy status, or future educational potential. Methodological research on large-scale assessment data has emphasized the importance of response quality and measurement characteristics, reinforcing the need to avoid attributing performance classifications to individual deficiencies or institutional failures without supporting evidence (Wang, 2025; Marcq & Braeken, 2025). The observed distribution demonstrates that minimum-proficiency attainment constitutes a substantial challenge in Indonesia across mathematics, reading, and science, while its descriptive nature prevents conclusions about the causal mechanisms underlying these outcomes. Further interpretation requires connecting these proficiency patterns to the broader implications of educational improvement, instructional support, and evidence-based policy evaluation.

Interpretation of Learning Decline and Implications for Educational Improvement

The longitudinal patterns of Indonesia's PISA achievement indicate that educational improvement requires consideration of both changes in national mean performance and the distribution of students attaining foundational competencies. The decline in reading achievement between 2012 and 2022, the reduction in mathematics performance following its 2015 peak, and the fluctuations in science achievement demonstrate that educational deterioration followed different trajectories across competency domains. These findings support a differentiated interpretation of learning decline rather than a uniform explanation based on a single assessment period or educational disruption. Research examining educational inequality and institutional learning conditions has similarly emphasized that national performance indicators require interpretation alongside differences in educational opportunities and school environments (Aditia & Széll, 2025; Steinmann & Olsen, 2022).

The temporal characteristics of the assessment results provide an important distinction between persistent educational weaknesses and disruptions associated with the COVID-19 pandemic. Indonesia's reading score declined from 397 in 2015 to 371 in 2018, establishing that substantial deterioration occurred before the widespread interruption of classroom instruction, while the subsequent decline to 359 in 2022 reflected a further reduction during an interval that included pandemic-related educational disruption. International evidence has documented learning losses associated with the pandemic, although the magnitude of deterioration varies across educational systems, student populations, and assessment conditions (Betthäuser et al., 2023; Castellano et al., 2025). The observed Indonesian trajectory cannot isolate the contribution of school closures from pre-existing educational conditions because the study uses aggregate national statistics without counterfactual comparisons or explanatory covariates.

Instructional improvement priorities can be informed by the differences between competency domains, particularly when the objective is to address foundational learning rather than raise national rankings alone. The PISA 2022 proficiency indicators demonstrate that mathematics had the highest proportion of students below Level 2, followed by reading and science, indicating that deficiencies in foundational competencies extended across the assessed curriculum. Research on mathematical learning strategies and scientific instruction suggests that cognitive demands and the organization of teaching activities are relevant considerations when designing learning support, although their effectiveness cannot be inferred directly from national assessment averages (Kim et al., 2024; Cairns, 2025). The descriptive results support the identification of mathematics and reading as areas requiring

particular policy attention, while intervention selection must remain conditional on evidence obtained through appropriately designed educational evaluations.

The practical interpretation of the findings requires separating descriptive evidence from instructional recommendations and causal explanations. Table 3 consolidates the observed achievement differences across consecutive assessment periods and the complete 2012–2022 interval, allowing the timing and magnitude of performance changes to be compared without introducing unobserved data. The calculations provide a basis for evaluating whether achievement deterioration was concentrated in particular assessment intervals or occurred consistently across domains. This approach is consistent with methodological research emphasizing careful interpretation of large-scale assessment indicators and the importance of distinguishing observed score differences from educational mechanisms that require additional empirical testing (Marcq & Braeken, 2025; Wang, 2025).

Table 3. Changes in Indonesia's PISA Achievement Scores Across Assessment Periods

Assessment period	Mathematics (Δ)	Reading (Δ)	Science (Δ)
2012–2015	+11	+1	+21
2015–2018	−7	−26	−7
2018–2022	−13	−12	−13
2012–2022	−9	−37	+1

Source: Calculated from published OECD PISA national mean scores for Indonesia, 2012–2022.

The comparative differences presented in Table 3 demonstrate that the direction of achievement change varied considerably across assessment intervals. Mathematics and science recorded their largest observed increases between 2012 and 2015, whereas reading experienced its largest interval decline between 2015 and 2018, indicating that the timing of performance deterioration differed across competency domains. The 2018–2022 interval produced reductions in all three subjects, although the preceding assessment patterns prevent those changes from being interpreted as evidence of a common causal mechanism. Comparative research on educational policy and learning culture supports attention to differences in educational environments when interpreting achievement outcomes, but the present analysis does not estimate the contribution of particular institutional arrangements (Purnasih et al., 2026).

The proportion of students below minimum proficiency provides a complementary basis for identifying instructional priorities that may not be apparent from mean achievement changes alone. In 2022, 82% of Indonesian students remained below Level 2 in mathematics, 75% in reading, and 66% in science, suggesting that foundational learning difficulties represented a substantial challenge across the assessed domains. Research examining intensive supplemental instruction in reading and mathematics has demonstrated the relevance of evaluating structured learning support over time, while studies of teacher feedback and reading self-efficacy highlight instructional processes that may warrant further investigation (Lindström-Sandahl et al., 2025; Wang et al., 2025). These findings provide a research-informed basis for considering targeted remediation and formative assessment, although no such instructional intervention was evaluated in the present study.

Educational recovery strategies must also consider differences in learning opportunities because national achievement indicators can conceal inequalities associated with household resources, institutional conditions, and instructional provision. Research investigating socioeconomic differences in mathematical and scientific literacy has identified educational disadvantage as a relevant dimension of achievement variation, while comparative studies have demonstrated the importance of examining inequality beyond national mean performance (Kastorff & Heine, 2025; Özdoğan Özbal, 2026; Maheshwari & Koria, 2025). Indonesian research on school governance and educational achievement similarly suggests that institutional conditions warrant attention when interpreting performance disparities (Miftakhi & Waluyo, 2025). The aggregate observations used in this investigation cannot determine whether achievement deterioration was concentrated among particular socioeconomic groups, making distributional analysis using student-level and school-level datasets a necessary direction for further research.

Teacher capacity and the continuity of instructional support represent additional considerations for interpreting educational recovery needs. Indonesian research has examined relationships involving digital literacy, learning arrangements, and educational difficulties during the pandemic period, while international investigations have considered the relevance of teacher support and socio-emotional conditions to learning outcomes (Safrizal et al., 2023; Fatayan et al., 2023; Golle et al., 2026; Lichand et al., 2024). These studies suggest that educational responses may need to address both academic competencies and the learning conditions under which those competencies develop. The present findings do not establish that differences in teacher capacity, digital resources, or socio-emotional support produced the observed national score changes, but they identify dimensions that could be investigated in subsequent explanatory and intervention-based studies.

Interpretation of PISA trends also requires recognition of the methodological limitations associated with comparing successive national assessment populations. The observed scores represent estimates of performance among sampled 15-year-old students in different assessment cycles rather than repeated measurements of identical individuals, preventing conclusions about individual learning progression. Research on large-scale assessment quality indicates that response behavior and measurement characteristics can influence the interpretation of achievement indicators, reinforcing the importance of transparent reporting and caution in drawing substantive conclusions from aggregate comparisons (Wang, 2025; Appels et al., 2024). The absence of individual-level observations, standard errors, and relevant explanatory variables limits the analysis to descriptive differences and prevents statistical testing of policy effects, socioeconomic mechanisms, or the independent contribution of pandemic-related disruptions.

The implications of the findings extend to the design of future educational evaluation and the prioritization of foundational learning within Indonesia's education system. The pronounced decline in reading achievement, the reversal of earlier mathematics improvements, and the substantial proportions of students below minimum proficiency indicate that national educational progress should be assessed through both achievement trajectories and competency distributions. Research on reading strategies, educational inequality, and instructional quality provides potential directions for subsequent investigation, particularly studies connecting school resources, student characteristics, and pedagogical practices with measurable achievement outcomes (Ghimire & Mokhtari, 2025; Yang, 2026; Patsis et al., 2026). A stronger empirical research agenda would combine nationally representative student- and school-level information with longitudinal or intervention-based designs capable of evaluating explanatory relationships rather than inferring them from aggregate score changes. The present contribution remains a reproducible descriptive account of Indonesia's PISA performance that clarifies the temporal characteristics of learning decline and identifies minimum proficiency as a central consideration for educational improvement.

CONCLUSION

The longitudinal analysis of Indonesia's PISA performance from 2012 to 2022 revealed distinct achievement trajectories across mathematics, reading, and science. Reading experienced the largest decline, decreasing by 37 points from 396 to 359, while mathematics declined by nine points from 375 to 366. Science exhibited a marginal increase of one point between the initial and final assessments, although its performance declined substantially after reaching its highest observed score in 2015. The assessment intervals demonstrated that educational deterioration was not uniform across competency domains, with reading showing a pronounced decline before the COVID-19 pandemic. Minimum-proficiency indicators further revealed that only 18% of Indonesian students attained Level 2 or higher in mathematics, 25% in reading, and 34% in science in 2022. These findings establish that Indonesia's educational challenges involve both declining achievement in particular domains and substantial proportions of students performing below foundational competency thresholds. The study contributes a reproducible descriptive assessment that integrates longitudinal achievement changes with minimum-proficiency distributions, offering a more differentiated interpretation of national educational performance than comparisons based exclusively on average scores or international rankings.

The findings support prioritizing foundational mathematics and reading competencies, strengthening targeted instructional assistance, and developing educational evaluation practices that distinguish achievement improvement from minimum-proficiency attainment. Educational policies

should consider differences across competency domains while directing attention toward students who have not demonstrated the competencies represented by internationally established proficiency thresholds. Formative assessment, structured remediation, and differentiated learning support constitute relevant areas for subsequent intervention evaluation, although their effectiveness was not directly examined in this study. The analysis is limited by its reliance on published aggregate PISA indicators, the absence of individual-level and school-level explanatory variables, and the inability to establish statistical significance or causal relationships between educational conditions and achievement changes. Differences between assessment populations also prevent the observed national trends from being interpreted as longitudinal changes among identical students. Future investigations should incorporate disaggregated achievement data, socioeconomic characteristics, instructional conditions, and appropriate explanatory or intervention-based designs to identify the mechanisms associated with persistent proficiency deficits. The principal implication is that educational progress in Indonesia should be evaluated not only through changes in national mean achievement but also through measurable improvements in the proportion of students acquiring foundational mathematical, reading, and scientific competencies.

REFERENCES

Acun Çelik, S., Özkan Elgün, İ., & Kalelioğlu, F. (2024). Assessment of student ICT competence according to mathematics, science, and reading literacy: Evidence from PISA 2018. Large-scale Assessments in Education, 12, Article 30. https://doi.org/10.1186/s40536-024-00218-7

Aditia, R., & Széll, K. (2025). Belonging matters: How context and inequalities shape student achievement in Indonesia. International Journal of Educational Research Open, 9, Article 100512. https://doi.org/10.1016/j.ijedro.2025.100512

Appels, L., De Maeyer, S., & Van Petegem, P. (2024). Re-thinking equity: The need for a multidimensional approach in evaluating educational equity through TIMSS data. Large-scale Assessments in Education, 12, Article 38. https://doi.org/10.1186/s40536-024-00227-6

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nature Human Behaviour, 7, 375–385. https://doi.org/10.1038/s41562-022-01506-4

Cairns, D. (2025). Balancing teacher-led and student-led learning in science: The importance of cognitive load. Large-scale Assessments in Education, 13, Article 27. https://doi.org/10.1186/s40536-025-00263-w

Castellano, K. E., McCaffrey, D. F., Xi, N., Jia, Y., & Hamilton, L. S. (2025). Exploring differential change in student performance during the COVID-19 pandemic across the United States. Large-scale Assessments in Education, 13, Article 11. https://doi.org/10.1186/s40536-025-00242-1

Cito, L., & Marôco, J. (2026). Beyond the average: Mapping educational inequality in Brazil with PISA 2022. Large-scale Assessments in Education, 14, Article 30. https://doi.org/10.1186/s40536-026-00302-0

Darmawan, I. G. N., & Dharmapatni, A. A. S. S. K. (2024). A multilevel analysis of student and school characteristics associated with 15-year-olds' reading performances: A Southeast Asian perspective. Large-scale Assessments in Education, 12, Article 40. https://doi.org/10.1186/s40536-024-00231-w

Fatayan, A., Ayu, S., Ghani, A. R. A., Kowiyah, & Azhar, N. C. (2023). The dynamics of learning loss for elementary students Jakarta in the new normal. Journal of Higher Education Theory and Practice, 23(6), 196–204. https://doi.org/10.33423/jhetp.v23i6.5967

Ghimire, N., & Mokhtari, K. (2025). Evaluating the predictive power of metacognitive reading strategies across diverse educational contexts. Large-scale Assessments in Education, 13, Article 4. https://doi.org/10.1186/s40536-025-00240-3

Golle, J., Lösche, P., Beißert, H., Wieland, A.-K., Rieber, R., Schult, J., Hasselhorn, M., Nagengast, B., & Trautwein, U. (2026). COVID-19-related school closures and students' achievement and motivation: Did teacher support make a difference. npj Science of Learning, 11, Article 48. https://doi.org/10.1038/s41539-026-00439-1

Kastorff, T., & Heine, J. H. (2025). Socioeconomic status and scientific literacy: Expanding educational inequalities during the COVID-19 pandemic—Insights from PISA 2018 and 2022.

Large-scale Assessments in Education, 13, Article 36. https://doi.org/10.1186/s40536-025-00273-8

Kastorff, T., Moser, S., Heine, J. H., & Kauertz, A. (2026). Global competence behavior: Exploring the relevance of students' scientific literacy, related attitudes, and values—Evidence from PISA 2018 across 52 countries. Large-scale Assessments in Education, 14, Article 3. https://doi.org/10.1186/s40536-025-00278-3

Kim, J., Chen, C.-W., & Wu, Y.-J. (2024). Exploration of the linear and nonlinear relationships between learning strategies and mathematics achievement in South Korea using the nominal response model: PISA 2012. Large-scale Assessments in Education, 12, Article 11. https://doi.org/10.1186/s40536-024-00198-8

León, J., & Martínez-Abad, F. (2025). The effect of grade retention on performance in Spanish students: A propensity score matching approach. Large-scale Assessments in Education, 13, Article 8. https://doi.org/10.1186/s40536-025-00243-0

Lichand, G., Christen, J., & Egeraat, E. V. (2024). Neglecting students' socio-emotional skills magnified learning losses during the pandemic. npj Science of Learning, 9, Article 28. https://doi.org/10.1038/s41539-024-00235-9

Lindström-Sandahl, H., Danielsson, H., Elwér, Å., Samuelsson, J., & Samuelsson, S. (2025). One-year follow-up of two intensive supplemental reading and mathematics programs in Swedish elementary school. Education Sciences, 15(12), Article 1678. https://doi.org/10.3390/educsci15121678

Maheshwari, A., & Koria, A. (2025). Educational inequality: A country-level comparison between OECD and non-OECD countries. Policy Futures in Education, 23(7). https://doi.org/10.1177/14782103251328399

Marcq, K., & Braeken, J. (2025). From framework to functionality: A cross-country analysis of PISA 2018 reading assessment framework's item features as determinants of item difficulty. Large-scale Assessments in Education, 13, Article 26. https://doi.org/10.1186/s40536-025-00261-y

Miftakhi, D. R., & Waluyo, M. E. (2025). Low gradients, low learning: School governance and the equity paradox in Indonesia's PISA 2022. Society, 13(3), 1320–1337. https://doi.org/10.33019/society.v13i3.921

Organisation for Economic Co-operation and Development. (2019). PISA 2018 results (Volume I): What students know and can do. https://doi.org/10.1787/5f07c754-en

Organisation for Economic Co-operation and Development. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. https://doi.org/10.1787/53f23881-en

Organisation for Economic Co-operation and Development. (2023). PISA 2022 results: Indonesia country note. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/indonesia_c2e1ae0e-en.html

Organisation for Economic Co-operation and Development. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Özdoğan Özbal, E. (2026). Digital use matters more than access: Socioeconomic inequality and mathematics achievement in the context of sustainable education (PISA 2022). Sustainability, 18(16), Article 8224. https://doi.org/10.3390/su18168224

Patsis, P., Yang Hansen, K., & Klapp, A. (2026). The role of instructional quality in explaining socioeconomic disparities in mathematics achievement: A multilevel structural equation modeling of Swedish TIMSS 2019 data. Large-scale Assessments in Education, 14, Article 34. https://doi.org/10.1186/s40536-026-00307-9

Purnasih, U., Subasman, I., Nuraeni, I., AS, S., Dirno, Khilal, R., & Amalia, H. (2026). Explaining PISA 2022 achievement disparities through educational policy and learning culture: A macro-comparative analysis of Indonesia, Brunei Darussalam, Portugal, and New Zealand. Journal of Education Research, 7(3), 1644–1659. https://doi.org/10.37985/jer.v7i3.3983

Safrizal, S., Yulnetri, Y., & Wulandari, N. (2023). Hubungan kemampuan literasi digital guru dengan learning loss pada siswa sekolah dasar di Nagari Lima Kaum. Jurnal Didaktika Pendidikan Dasar, 7(3), 829–842. https://doi.org/10.26811/didaktika.v7i3.1147

Steinmann, I., & Olsen, R. V. (2022). Equal opportunities for all? Analyzing within-country variation in school effectiveness. Large-scale Assessments in Education, 10, Article 2. https://doi.org/10.1186/s40536-022-00120-0

Wang, M. (2025). Data quality disparities in large-scale assessments: Insufficient effort responding across student groups, schools, and cultures. Large-scale Assessments in Education, 13, Article 21. https://doi.org/10.1186/s40536-025-00260-z

Wang, X., Zhang, Q., & Davis, M. H. (2025). Positive teacher feedback and adolescents' reading self-efficacy: A quasi-experimental analysis using PISA 2018. Large-scale Assessments in Education, 13, Article 17. https://doi.org/10.1186/s40536-025-00253-y

Wibowo, A. H. (2022). The impact of learning from home (LFH) during COVID-19 pandemic on students' learning loss. Journal of English Teaching, Literature, and Applied Linguistics, 6(2), 82–87. https://doi.org/10.30587/jetlal.v6i2.4161

Yang, R. (2026). Socioeconomic inequalities in reading literacy and learning-related characteristics: Evidence from PISA 2022. Journal of Education Research, 7(3). https://doi.org/10.37985/jer.v7i3.3480