International Journal of Education, Humanities and Social Sciences

Article

Why Has PISA Become a Leading Framework for Competency-Based Assessment?

Damu Wu [1] and Altangerel Balgan [1,*]

[1] Department of Education Studies, The National University of Mongolia, Ulaanbaatar, Mongolia
* Correspondence: Altangerel Balgan, Department of Education Studies, The National University of Mongolia, Ulaanbaatar, Mongolia

Abstract: As international student assessment has evolved from curriculum-based knowledge testing to large-scale standardised comparison and, more recently, to competency-oriented evaluation, PISA has increasingly been regarded as a leading framework for competency-based assessment. Yet its representative status is often assumed rather than systematically explained. This article examines why PISA has become especially influential in contemporary educational evaluation and why it has come to symbolise a broader shift in assessment logic. Drawing on document analysis, comparative analysis, and theory-informed review, the article analyses the historical evolution of international student assessment, compares the distinctive orientations of TIMSS, PIRLS, and PISA, and interprets PISA through the lenses of competency-based assessment, educational equity, school effectiveness, and human capital and lifelong learning. The article finds that PISA became representative not simply because it is a large-scale international assessment, but because it redefined educational assessment in five important ways: by shifting the focus from knowledge acquisition to knowledge application; by embedding tasks in real-life or near-real-life contexts; by foregrounding transfer, problem solving, and cross-context performance; by operating through a framework not fully tied to any single national curriculum; and by linking student performance to broader questions of educational quality, equity, and policy learning. The article argues that PISA's significance lies not merely in international ranking, but in offering an internationally visible model of assessment aligned with competency, context, and development. At the same time, its policy relevance depends on careful local interpretation rather than direct policy borrowing.

Keywords: pisa; competency assessment; student assessment; educational evaluation; policy learning

Received: 14 February 2026
Revised: 31 March 2026
Accepted: 10 April 2026
Published: 13 April 2026

1. Introduction

Educational assessment has changed profoundly over the past several decades [1]. Earlier traditions of assessment were largely organized around curriculum content, textbook knowledge, and the measurement of mastery within specific subjects. Their core purpose was to determine whether students had learned what schools intended to teach, often through standardized items designed to capture recall, comprehension, and limited forms of procedural application. In this sense, assessment functioned primarily as a mechanism of achievement verification, selection, and sorting.

As national education systems expanded and governments increasingly sought comparable evidence on school performance, assessment entered a second phase characterized by large-scale standardized comparison. In this phase, the central concern was no longer limited to what individual students had learned in relation to a specific curriculum, but extended to how schools, regions, and countries compared under common measurement frameworks. Standardization, comparability, sampling, and monitoring became central features of assessment.

A further transformation occurred when assessment began to respond more directly to the demands of knowledge societies, globalization, technological change, and lifelong learning. In this context, the crucial question was no longer simply whether students had acquired curriculum knowledge, but whether they could mobilize knowledge and skills to address unfamiliar problems in complex and meaningful situations. It is within this broader transformation that the OECD's Programme for International Student Assessment (PISA) emerged as a particularly influential framework. OECD defines PISA as an assessment of the extent to which 15-year-old students nearing the end of compulsory education have acquired the knowledge and skills essential for full participation in modern societies, and it explicitly emphasizes students' ability to use reading, mathematics, and science knowledge and skills to meet real-life challenges [1].

This distinction is central to the argument of the present article. In many policy and academic discussions, PISA is described as influential because of its international reach, its role in public debate, or its impact on reform agendas. These claims are not wrong, but they are incomplete. They explain PISA's visibility, yet they do not fully explain why PISA, rather than other international assessments, has become a leading reference point in debates on competency-based assessment. TIMSS and PIRLS are also large-scale, cross-national assessments with established technical frameworks, yet they occupy a somewhat different place in international policy discourse [2]. TIMSS is designed to provide countries with information about students' mathematics and science achievement, while PIRLS focuses on reading literacy at the fourth grade and the conditions supporting children's reading development.

This article argues that PISA's representative status can only be fully understood when it is situated within the historical development of international student assessment and examined at the level of assessment logic. PISA is not merely one international testing program among others. Rather, it crystallizes a broader shift from curriculum-bound measurement to competency-oriented evaluation [2]. It retains the scale, standardization, and comparability of earlier large-scale assessments, while extending assessment beyond school-based achievement by foregrounding literacy, application, transfer, and contextual performance. This line of analysis matters because it moves the discussion away from simple visibility or ranking effects and toward a deeper question: what kind of educational achievement is each assessment framework designed to represent?

The article addresses four research questions [1]. First, how has international student assessment evolved from knowledge testing to competency-oriented evaluation? Second, in what ways does PISA differ from TIMSS and PIRLS at the level of assessment logic? Third, why has PISA become a leading framework for competency-based assessment? Fourth, what implications does this framework have for contemporary educational evaluation reform?

2. Methods

This article adopts a theory-informed review design and follows an IMRaD structure adapted for a conceptual study. It is not an empirical article based on original statistical analysis, experimental design, or primary fieldwork. Instead, it combines document analysis, comparative analysis, and theoretical interpretation [3]. This methodological choice is appropriate because the question addressed here—why PISA has become a leading framework for competency-based assessment—is explanatory and conceptual rather than causal in a narrowly empirical sense. The study therefore focuses on the purposes, frameworks, and policy uses of major international assessments rather than on hypothesis testing with newly generated data.

The first component of the method is document analysis. The study reviews official framework documents and related policy materials from the principal international assessment programmes discussed in the article. The core documentary sources are the OECD PISA frameworks and the official TIMSS and PIRLS assessment frameworks produced by the TIMSS & PIRLS International Study Center at Boston College. These sources define the target populations, purposes, assessment domains, and framework

logics of the programmes being compared and therefore provide the most authoritative basis for analysis. OECD's PISA frameworks present literacy as an applied construct concerned with the use of knowledge and skills in meaningful contexts, whereas TIMSS and PIRLS remain more closely tied to subject achievement and reading development within schooling [2].

The second component is comparative analysis. The article compares TIMSS, PIRLS, and PISA across five analytical dimensions: target population, content orientation, task context, conception of competence, and policy function. These dimensions are not arbitrary; they are derived from the core differences among the three programmes, especially the extent to which they are tied to school curricula, developmental reading, or real-world literacy and application. The comparison therefore moves beyond surface contrasts and identifies the deeper differences in assessment logic that shape each programme's significance [4].

The third component is theoretical interpretation [4]. To explain why PISA has become especially influential, the article draws on four theoretical perspectives: competency-based assessment, educational equity, school effectiveness, and human capital and lifelong learning. Competency-based assessment helps explain PISA's shift from knowledge possession to knowledge use. Educational equity clarifies why PISA's contextual and background data are important for interpreting differences in student performance. School effectiveness highlights the relevance of school- and system-level conditions, while human capital and lifelong learning perspectives explain why literacy, transfer, and problem solving resonate so strongly with contemporary educational goals and policy expectations. These perspectives are not treated as loosely related background ideas; they serve as an interpretive framework for understanding PISA's representative status.

The scope of the article is deliberately limited [4]. It does not attempt to provide a statistical comparison of countries or a field-based evaluation of national reforms. Instead, it focuses on explaining conceptually why PISA has become a leading framework for competency-based assessment. This narrower scope strengthens the coherence of the argument and makes the article better suited to journals in comparative education, educational assessment, and education policy.

3. Results

3.1. This article finds that international student assessment has undergone a paradigm shift from knowledge testing to competency-oriented evaluation.

The first finding is historical and foundational. International student assessment did not begin as competency-based evaluation. It developed through a sequence of overlapping but distinguishable phases: an early phase centered on subject knowledge testing, a later phase focused on standardized large-scale comparison, and a more recent phase centered on competency and literacy [1]. In the first phase, success was defined mainly in terms of curriculum mastery and accurate reproduction of school knowledge. In the second phase, comparability and monitoring became more important, and assessment gained greater value for system governance. In the third phase, the focus shifted toward whether students could use knowledge and skills in meaningful situations beyond the classroom.

This historical sequence matters because it changes the meaning of educational achievement. Under a knowledge-testing model, achievement is usually interpreted as mastery of prescribed content. Under a competency-oriented model, achievement includes the capacity to mobilize knowledge and skills in contexts that are not reducible to the curriculum itself [4]. PISA occupies a central place in this transition because it inherits the standardization and comparability of large-scale assessment while extending its focus toward application, transfer, and contextual performance. OECD's own description of PISA supports this interpretation by defining the program in terms of

students' ability to use knowledge and skills to meet real-life challenges rather than simply to reproduce what has been taught.

This historical account supports a three-stage interpretation of international student assessment: traditional knowledge testing, standardized comparison, and competency-oriented evaluation. Within this sequence, PISA can be understood as the most influential expression of the third stage, because it extends large-scale assessment beyond curriculum mastery toward literacy, transfer, and contextualized performance [5]. That is why PISA has come to symbolize not only a program, but a broader shift in assessment thinking.

3.2. This article finds that PISA differs from TIMSS and PIRLS not mainly by subject coverage, but by assessment logic.

The second finding is comparative. TIMSS, PIRLS, and PISA are often grouped together as major international assessments, yet their differences are deeper than domain labels or target ages. The key distinction lies in what each program considers to be the meaning of educational performance. TIMSS is organized around mathematics and science achievement in schooling. PIRLS focuses on reading literacy at the fourth grade during a foundational phase of learning. PISA, by contrast, defines reading, mathematics, and science as forms of literacy linked to the use of knowledge and skills in real-life contexts [6].

This means that TIMSS and PIRLS remain more closely tied to school-based learning trajectories, whereas PISA moves more directly toward broader capability, future participation, and social relevance. The difference is not merely technical [7]. It reflects different assumptions about what counts as important educational achievement. TIMSS asks, in effect, how well students are learning school mathematics and science. PIRLS asks how students are developing reading literacy during a key stage of primary schooling. PISA asks whether students nearing the end of compulsory education can use knowledge and skills in ways relevant to life beyond the classroom. A comparative analysis of TIMSS, PIRLS, and PISA will be presented in Table 1.

Table 1. Comparative Analysis of TIMSS, PIRLS, and PISA

Comparative Dimension	TIMSS	PIRLS	PISA
Target population	Primarily assesses students in Grades 4 and 8, with a focus on mathematics and science achievement	Primarily assesses Grade 4 students, focusing on a key stage in children's reading development	Primarily assesses 15-year-old students; sampling is age-based rather than grade-based, with a stronger international comparative orientation
Assessment orientation	Emphasises mastery of curriculum content and instructional effectiveness; largely curriculum-based and subject-oriented	Emphasises reading comprehension and its foundational role in subsequent learning; oriented toward basic learning capacity	Emphasises knowledge transfer, real-life application, and problem solving; more clearly reflects competency- and literacy-oriented assessment
Content framework	Organised around mathematics and	Organised around literary and	Organised around reading, mathematical,

	science content domains and cognitive domains, with emphasis on disciplinary knowledge structures	informational texts and different comprehension processes, focusing on reading comprehension	and scientific literacy, with emphasis on the mobilisation of knowledge across contexts and integrated judgement
Relationship to curriculum	Most closely aligned with school curricula, with emphasis on the coherence of the intended, implemented, and attained curriculum	Closely connected to elementary reading instruction and language learning in basic education, serving the monitoring of foundational skills	Not fully tied to any single national curriculum; instead emphasises key competencies widely required in modern society
Policy use	Better suited to evaluating subject teaching quality, curriculum implementation, and school achievement	Better suited to monitoring early reading development, the quality of basic education, and learning support conditions	Better suited to curriculum reform, policy adjustment, and cross-national comparisons of educational quality
Analytical value for educational equity	Can reveal differences in subject achievement and curriculum implementation across schools and regions, but remains primarily centred on academic outcomes	Can be combined with background variables such as family reading environment and school resources to analyse disparities in early educational opportunities	More conducive to examining fairness in educational opportunities, outcomes, and competency development by linking socioeconomic background, resource allocation, and student performance

From the perspective of target population, TIMSS and PIRLS are both anchored in specific stages of basic education [5]. TIMSS focuses on mathematics and science achievement among students in Grades 4 and 8, while PIRLS concentrates on the reading literacy of Grade 4 students, a phase widely regarded as critical in the transition from "learning to read" to "reading to learn." Both assessments therefore place stronger emphasis on stage-specific learning outcomes within formal schooling. PISA, by contrast, targets 15-year-old students and adopts an age-based rather than grade-based sampling logic. As a result, it is less concerned with performance at a particular grade level and more concerned with whether students nearing the end of compulsory education have acquired the foundational capabilities needed for future participation in society.

In terms of assessment orientation, TIMSS and PIRLS remain closer to the logic of large-scale standardized comparison. Both are designed to compare student performance across countries and regions within a common framework, although they differ in substantive emphasis. TIMSS is more explicitly curriculum- and subject-oriented, with a stronger focus on the extent to which students have achieved expected learning outcomes in mathematics and science. PIRLS, in contrast, places greater weight on foundational reading comprehension and the conditions that support its development. PISA differs from both in that it more clearly embodies a competency- and literacy-oriented approach. Its central concern is not simply how much students have learned, but whether they can mobilize knowledge, interpret information, and make effective judgments and responses in real or near-real-life situations. In this sense, the contrast among the three assessments is consistent with the broader shift in international educational evaluation from standardized comparison toward competency-oriented assessment.

Differences also become evident in the content frameworks and their relationship to curriculum. TIMSS is closely aligned with school mathematics and science curricula and employs the analytical distinction among the intended, implemented, and attained curriculum to examine the relationship between instruction and outcomes. PIRLS, although broader than simple word recognition or decoding, remains firmly centered on reading comprehension as a foundational learning capacity in primary education and therefore retains a strong connection to classroom instruction and basic education curricula. PISA, by contrast, is not fully dependent on the curriculum standards of any single country. Instead, it constructs its framework around reading, mathematical, and scientific literacy as competencies broadly required in modern society. This gives greater prominence to knowledge transfer, cross-context application, and integrated problem solving than to curriculum mastery alone.

The three assessments also differ in their policy functions. TIMSS and PIRLS are especially useful for analyzing curriculum implementation, disciplinary foundations, and the quality of basic education. TIMSS provides evidence on the extent to which mathematics and science curriculum goals have been achieved, on the effectiveness of instructional implementation, and on differences in subject teaching across educational systems. PIRLS is particularly valuable for identifying the foundational level of children's reading development and, when combined with background variables from families and schools, for evaluating the effectiveness of early learning support systems. PISA, by virtue of its literacy-oriented, cross-curricular, and cross-national design, is more readily used in discussions of curriculum reform, policy evaluation, and international educational competitiveness.

A further distinction lies in the analytical value of these assessments for the study of educational equity. All three can reveal disparities among groups, schools, and regions through background questionnaires and large-scale data analysis, but they do so with different emphases. TIMSS and PIRLS are particularly useful for examining how curriculum implementation, school resources, and family learning support are associated with students' academic performance, making them well suited to analyzing disparities in educational conditions and outcomes during the basic education stage. PISA, however, is methodologically more advantageous for examining equity in relation to competency development because it places greater emphasis on performance in real-life contexts and more readily connects socioeconomic background, educational opportunity, and literacy outcomes. This makes it especially valuable for cross-national research on both educational quality and educational equity.

Overall, TIMSS and PIRLS are better suited to analyzing curriculum implementation, subject-based foundations, and the quality of teaching and learning in basic education, whereas PISA is better suited to cross-national comparison, competency analysis, and research on educational equity. For the purposes of the present study, a comparison between China and Mongolia requires not only an examination of differences in academic performance, but also an analysis of disparities in core competencies, real-world problem-solving capacity, and the distribution of educational opportunity. In this respect, PISA

provides the most appropriate analytical instrument. Its use is consistent both with the broader movement in international educational assessment from knowledge-based to competency-oriented evaluation and with the present study's concern for educational quality and equity.

3.3. This article finds that PISA became representative because it shifted the object of assessment from knowledge possession to knowledge use.

The third and most important finding is conceptual. Under traditional curriculum-based assessment, the focus is often on students' possession of knowledge: whether they know facts, remember procedures, and reproduce learned material. Under PISA, the focus shifts to knowledge use: whether students can mobilize what they know and can do in response to meaningful and unfamiliar challenges.

This shift changes the criteria of success. Success is no longer defined mainly by recall and procedural completion, but by the ability to interpret, reason, apply, and respond under conditions that are less predictable than classroom routines. This is precisely why PISA has become central to debates on competency-based assessment. It does not simply use the language of literacy; it operationalizes a broader educational aspiration that students should be able to use what they know in contexts that matter. OECD's PISA frameworks consistently define reading, mathematical, and scientific literacy as applied constructs that involve interpretation, reasoning, and use rather than mere content reproduction.

This interpretation is further strengthened by research linking international assessment to broader debates on knowledge economies and lifelong learning [8]. PISA's emphasis on reading, mathematics, and science literacy in real-world situations aligns closely with the capacities increasingly regarded as important in knowledge-based societies. These are not only academic abilities; they are capabilities associated with future participation, adaptability, and sustained development. This alignment strengthens PISA's status as a leading framework because it links assessment to broader societal expectations rather than to school achievement alone.

3.4. This article finds that PISA's representative status also depends on contextualised task design and cross-context transfer.

A fourth finding concerns how PISA operationalizes competency. Many assessment systems refer to competencies, skills, or core literacies, but not all of them translate those ideas into assessment design. PISA became especially influential because it embedded competency in task structures oriented toward context, interpretation, transfer, and problem-solving. Its tasks are frequently situated in personal, public, occupational, or educational settings, requiring students to use knowledge in ways closer to everyday life than conventional school tests do. OECD's assessment and analytical frameworks specify contexts, processes, and content together, which is one reason PISA has become central to discussions of applied literacy.

This feature matters because competency-based assessment is not defined by rhetoric alone. It depends on whether tasks actually require the coordinated use of cognitive resources in context. Traditional achievement tests may include items labeled "application" or "reasoning," yet these often remain confined within familiar subject routines. PISA goes further by asking students to interpret information embedded in wider contexts, which makes the assessment less dependent on direct curriculum recall and more dependent on transfer and adaptive use.

PISA tasks are situated in real or near-real contexts, and the program's reading, mathematics, and science constructs all imply the integration of knowledge, skills, strategies, and judgment. Because real-life problems do not present themselves neatly within disciplinary boundaries, PISA's literacy framework is especially well suited to discussions of cross-disciplinary and transferable competence. This helps explain why PISA has often been taken as a reference point in national debates on key competencies and core literacies.

3.5. This article finds that PISA became representative because its framework is not fully tied to any single national curriculum.

A fifth finding is that PISA's representative status depends partly on its relative independence from any one country's curriculum structure. This does not mean that PISA is detached from content or entirely culture-free. Rather, it is designed around a comparatively general framework of literacy and application intended to be meaningful across systems. PISA samples students by age rather than grade and defines domains through literacy constructs linked to participation in modern society [1, 9].

This design choice matters because when assessment is tightly tied to specific curriculum sequences, textbooks, or instructional pacing, cross-national comparability becomes more difficult to interpret. A high score may reflect curriculum alignment rather than broader capability; a low score may reflect timing or exposure rather than limited ability. By contrast, PISA's age-based sampling and literacy-oriented design make it easier to frame results as indicators of broader capability rather than as reflections of curriculum timing alone.

This relative independence from single-system curricular design therefore enhances both the scope and the policy relevance of PISA. TIMSS and PIRLS are highly valuable for monitoring achievement and development within schooling, but PISA's broader framing makes it especially resonant in debates about future capability, system preparedness, and the social relevance of educational outcomes.

3.6. This article finds that PISA gained special policy relevance because it links performance to educational quality, equity, and system improvement.

The sixth finding is that PISA's influence cannot be explained by assessment design alone. PISA also became representative because its results are embedded in a broad analytical framework linking student performance to issues of educational quality, equity, school conditions, and policy learning. Through contextual questionnaires and system-level indicators, PISA can be used to ask not only how well students perform, but also how performance is distributed, what background factors are associated with it, and what institutional conditions may matter for improvement. OECD's technical and analytical documentation explicitly positions PISA as an instrument for examining not only performance, but also the contextual factors associated with learning and equity.

This broader interpretive capacity makes PISA more than a test. It becomes a tool through which systems can diagnose quality and inequality together. One reason for this is that PISA has consistently encouraged the interpretation of results through background variables, including socioeconomic conditions and school environments [10]. That analytical design aligns strongly with educational equity and school-effectiveness perspectives.

PISA's policy relevance is also well documented in the literature on international benchmarking and educational governance [11]. It describes how national policy actors use PISA to evaluate and improve school-system performance, while it is argued that PISA has become influential as a technology of governing education "by numbers." Further analysis shows that the expansion of PISA has strengthened new global modes of governance in education, and an analysis of Germany after the "PISA shock" illustrates how results can become catalysts for reform discourse. Together, these studies explain why PISA has often generated stronger public and political response than other assessment programmes. It not only reports outcomes; it also helps define what should be seen as a systemic educational problem.

4. Discussion

The findings of this article suggest that PISA's representative status cannot be understood through international diffusion alone. PISA did not become a leading framework for competency-based assessment merely because it expanded globally or because its results became highly visible [12]. It became representative because it aligned with, and helped consolidate, a deeper shift in the logic of educational evaluation. This

shift involved a movement away from curriculum-bound knowledge measurement toward forms of assessment that emphasize application, transfer, context, and future-oriented capability. In this sense, PISA is best understood less as a stand-alone program than as a crystallization of a broader transformation in assessment thinking.

The article makes three principal contributions. First, it offers a historical explanation of PISA's prominence. Much writing on PISA discusses its global visibility, league-table effects, or policy consequences. Those discussions are useful, but they often begin after PISA has already been treated as important. The present article asks a prior question: why did PISA come to be treated as a central reference point in the first place? The answer proposed here is that PISA emerged at a moment when the purposes of assessment were being redefined, and it gave institutional form to the idea that educational success should include the ability to use knowledge in meaningful situations.

Second, the article clarifies that the most important difference between PISA and other major international assessments lies in assessment logic rather than subject labels. TIMSS, PIRLS, and PISA all provide valuable forms of evidence, but they illuminate different aspects of educational performance. TIMSS remains closely tied to mathematics and science achievement in schooling, while PIRLS focuses on reading development at a foundational stage; PISA defines domains through literacy constructs linked to application and social participation. Recognizing this distinction prevents misleading comparisons and clarifies why PISA, in particular, became so influential in debates on competency-based reform.

Third, the article shows that PISA's influence rests on the conjunction of competency-oriented design and policy usability. PISA matters not only because it measures something different, but because it packages that measurement within a framework that policymakers can use to think about quality, equity, and system performance. The inclusion of contextual questionnaires and system-relevant indicators has made PISA especially visible in reform discourse. This helps explain why PISA has often shaped public discussion and policy agendas more forcefully than many other assessment programs.

At the same time, the article has clear limits. It is a conceptual and review-based study rather than an empirical test of whether PISA produces better educational outcomes than other assessments or whether its influence is uniformly beneficial across settings. It relies primarily on official frameworks, policy documents, and analytical interpretation [13]. For that reason, it explains why PISA has become representative, but it does not claim that PISA is normatively superior in every dimension or equally appropriate for every national context.

A second limit concerns transferability. Although PISA is often treated as a global reference point, its framework should not be mistaken for a universally applicable template for reform. The very features that make PISA analytically attractive—comparability, literacy constructs, contextualized tasks, and policy visibility—do not eliminate national differences in curriculum, language, educational culture, governance, or resource distribution. The most meaningful use of PISA therefore lies in interpretation and adaptation rather than mechanical borrowing [6, 14]. That conclusion is also consistent with the policy-impact literature, which shows that PISA's effects are mediated through national actors, institutions, and reform agendas rather than simply imposed from outside.

These limits lead directly to the article's implications. For research, international assessment should be studied not only through rankings and outcomes, but through the logics of measurement embedded in framework design [12]. For policy, PISA should be treated primarily as a diagnostic and reflective instrument rather than as a competitive scoreboard. For reform, the most valuable lesson from PISA is not the imitation of item formats or ranking practices, but the shift toward competency, context, and evidence-based diagnosis. Used in this way, PISA remains important not because it settles educational debates, but because it helps reframe them.

5. Conclusion

This article has argued that PISA became a leading framework for competency-based assessment because it reflects and operationalizes a broader transformation in the purposes and logic of educational evaluation. It emerged from a shift away from curriculum-bound knowledge testing toward forms of assessment that value application, transfer, contextual performance, and future-oriented capability. Unlike other assessments, which remain more closely associated with subject achievement and foundational reading development within schooling, PISA defines reading, mathematics, and science as forms of literacy linked to participation in modern society.

Its representative status, however, does not lie in visibility alone. PISA became influential because it combined a distinctive conception of knowledge use with contextualized task design, relative independence from any single national curriculum, large-scale comparability, and a broad analytical framework linking outcomes to quality, equity, and system improvement. These characteristics made PISA more than a testing program. They made it a visible model of what competency-based assessment can mean in practice.

The article also suggests that PISA's importance should not be reduced to international ranking or mechanical policy borrowing. Its deeper significance lies in the way it reframes educational achievement and invites systems to consider not only what students know, but what they can do with what they know in contexts that matter. Used in that sense, PISA remains important not because it closes educational debates, but because it helps redirect them.

References

1. I. V. Mullis and M. O. Martin, "TIMSS 2019 Assessment Frameworks," International Association for the Evaluation of Educational Achievement, Amsterdam, The Netherlands, 2017.
2. I. V. Mullis and M. O. Martin, "PIRLS 2021 Assessment Frameworks," International Association for the Evaluation of Educational Achievement, Amsterdam, The Netherlands, 2019.
3. L. De Bortoli, C. Underwood, and S. Thomson, "PISA 2022. Reporting Australia's results. Volume I: Student performance and equity in education," 2023.
4. J. Scheerens, "Educational effectiveness and ineffectiveness. A critical review of the knowledge base," 2016.
5. C. Sälzer and N. Roczen, "Assessing global competence in PISA 2018: Challenges and approaches to capturing a complex construct," International Journal of Development Education and Global Learning, vol. 10, no. 1, 2018.
6. B. Barbot and J. C. Kaufman, "PISA 2022 creative thinking assessment: Opportunities, challenges, and cautions," The Journal of Creative Behavior, vol. 59, no. 1, p. e70003, 2025.
7. S. Breakspear, "The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance," 2012.
8. Z. Özaydın and Ç. Arslan, "Assessment of mathematical reasoning competence in accordance with PISA 2021 mathematics framework," Journal of Theoretical Educational Science, vol. 15, no. 3, pp. 453-474, 2022.
9. S. Sellar and B. Lingard, "The OECD and global governance in education. Testing regimes, accountabilities and education policy," pp. 182-198, 2017.
10. E. A. Hanushek and L. Woessmann, "The knowledge capital of nations: Education and the economics of growth," MIT Press, 2023.
11. A. Z. Alfaruqi and N. Nurwahidah, "Reflection on Indonesia's PISA scores and the 2024 madrasah teacher competency assessment results: Challenges in enhancing teacher competence," Jurnal Pendidikan IPS, vol. 15, no. 1, pp. 11-19, 2025.
12. J. Amir, A. Dalle, S. Dj, and I. Irmawati, "PISA assessment on reading literacy competency: Evidence from students in urban, mountainous and island areas," Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran, dan Pembelajaran, vol. 9, no. 1, pp. 107-120, 2023.
13. A. S. Marleny, Zukardi, and R. I. I. Putri, "Systematic literature review: Development of PISA mathematics minimum competency assessment questions in tourism contexts," in AIP Conference Proceedings, vol. 3052, no. 1, p. 020076, AIP Publishing LLC, Apr. 2024.
14. A. Schleicher, "Can competencies assessed by PISA be considered the fundamental school knowledge 15-year-olds should possess?," Journal of Educational Change, vol. 8, no. 4, pp. 349-357, 2007.