

APPENDIX A

This appendix lists selected examples as mentioned by the education system in their encyclopedia chapter per category. The examples were selected based on relevance, level of concreteness, and geographic representation.

Data source for policymakers and educators	
TIMSS (50)	PIRLS (44)
TIMSS data were used by education researchers and policy bodies during the preparation of the 2014 national core curriculum, with widespread dissemination and debate on national performance. (<i>Finland</i>)	PIRLS results were actively disseminated to primary school teachers and used by the Coordination Center for Reading to inform teacher education programs and identify professional development topics, thus directly impacting educational policymaking and classroom practice. (<i>Austria</i>)
Policymakers and educators used TIMSS (and other studies) to analyze and revise mathematics and science curricula. (<i>England</i>)	Germany used PIRLS as part of its national education monitoring strategy, where results are fed into government policymaking and school-level instructional improvement, emphasizing evidence-based decision-making and practice. (<i>Germany</i>)
TIMSS data raised awareness among policymakers about structural weaknesses and curriculum misalignment. (<i>Romania</i>)	PIRLS findings are regularly shared with policymakers and educators in Singapore to guide the development of teaching strategies, reading resources, and curriculum improvements, especially in response to identified gaps in student reading behavior and multiliteracy skills. (<i>Singapore</i>)
Reviewed or revised curriculum	
TIMSS (49)	PIRLS (22)
TIMSS data has informed curriculum development by identifying areas where students struggle—such as specific mathematical concepts—prompting targeted adjustments to address these challenges. (<i>Hong Kong SAR</i>)	PIRLS participation led to revision of the Arabic language curriculum to strengthen students' reading competencies by making comprehension processes a central focus in the design of reading activities and textbooks. (<i>Bahrain</i>)
TIMSS results were used to revise primary and lower secondary mathematics and science curriculums in 2008 and again in 2022, leading to nationwide curriculum implementation and the integration of updated content across all grades by 2024. (<i>Lithuania</i>)	Based on PIRLS results the country formed committees to revise the Arabic language curriculum, developed new textbooks and teacher guides using PIRLS items, and embedded PIRLS findings into national education strategies. (<i>Jordan</i>)
TIMSS 2019 results were used to provide the Curriculum Department with the national report and accompanying recommendations to revise the mathematics and science curricula. (<i>UAE</i>)	Oman revised its Arabic language curriculum for Grades 1 to 12 based on PIRLS 2016 indicators, with a specific focus on enhancing reading comprehension skills from the early grades, including a targeted redevelopment of the Grades 1 to 4 curriculum between 2018 and 2020. (<i>Oman</i>)
Introduced or evaluated reforms or special support programs	
TIMSS (44)	PIRLS (37)
Armenia systematically adopted TIMSS methodologies to transform its entire assessment landscape, including continuous assessment, unified exams, and decisions on curriculum and textbook reforms. (<i>Armenia</i>)	Austria used PIRLS data explicitly to inform teacher education through the Coordination Center for Reading and to guide policy debates on educational standards, contributing to reforms such as standardized school leaving exams and a legislative decree on reading instruction aimed at better transitions between Kindergarten and primary school. (<i>Austria</i>)
Multiple administrative units in Bosnia & Herzegovina introduced or revised curricula, launched external assessments, and initiated extracurricular programs in response to TIMSS. They developed student achievement standards, conducted professional development, and created detailed policy reports and workshops to implement changes across regions. (<i>Bosnia & Herzegovina</i>)	Following the PIRLS 2016 results, Flanders launched a wide set of reforms including professional development programs, central testing policies, and the creation of a Reading Offensive strategy (2021–2030), all explicitly aimed at addressing reading comprehension issues highlighted by PIRLS. (<i>Belgium, Flemish</i>)
TIMSS insights led to the revision of national examination instruments (SPM) to include Higher-Order Thinking Skills (HOTS) and apply TIMSS-like cognitive domains. The reforms were explicitly timed to align with TIMSS cycles and aimed to improve analytical, evaluative, and creative reasoning in student assessments. (<i>Malaysia</i>)	After PIRLS 2001, Russia moved from oral reading assessments to written tests modeled on PIRLS at both the regional and federal levels. The PIRLS framework influenced the development of national programs like Reading. Working with a text, part of the Federal State Education Standards. (<i>Russian Federation</i>)

Initiate political discussions on national reforms	
TIMSS (16)	PIRLS (20)
TIMSS results prompted national debate over achievement gaps, student engagement, and teaching quality. (<i>Chinese Taipei</i>)	PIRLS participation led to nationwide discussions and direct curriculum reforms in Arabic language education, with PIRLS data explicitly used to guide the Education Strategic Plan 2018–2022 and the National Strategy for Human Resource Development, forming the basis for long-term national reform monitoring. (<i>Jordan</i>)
TIMSS 2019 results triggered a major public and political debate over declining mathematics scores. (<i>Denmark</i>)	Persistent declines in PIRLS results—especially noted in the 2016 cycle—prompted the Minister of National Education to publicly call for a “reading priority” in elementary schools, triggering reforms like smaller class sizes, a renewed focus on reading fundamentals, and new teacher training initiatives. (<i>France</i>)
Romania used TIMSS data to raise policymakers’ awareness of systemic challenges in education, prompting discussions on the need for curriculum and assessment reform. This led to the adoption of the 2023 Education Act, which introduced mandatory national assessments at the end of Grades 2, 4, and 6. (<i>Romania</i>)	After PIRLS 2016 revealed that 78 percent of children couldn’t read for meaning, this statistic was cited in the President’s State of the Nation address and used by the Human Rights Commission to petition the Constitutional Court to recognize basic reading as a constitutional right, placing PIRLS at the heart of a major political debate. (<i>South Africa</i>)
Implemented or revised national assessments	
TIMSS (11)	PIRLS (12)
Armenia used TIMSS methods and processes as a blueprint for creating a comprehensive national assessment system across multiple subjects (e.g., Armenian language, history, sciences, and foreign languages). TIMSS directly influenced the design and introduction of these assessments, which now include continuous, national, and unified exams, and inform secondary school graduation and university entrance exams. (<i>Armenia</i>)	Following participation in PIRLS, Azerbaijan implemented a centralized national assessment of reading and mathematics for fourth-grade students. The new assessment framework, publicly funded and widely disseminated, was explicitly designed to monitor education quality and address gaps identified through PIRLS data. (<i>Azerbaijan</i>)
TIMSS helped shape the development of Hungary’s National Assessment of Basic Competencies (NABC) starting in 2001. TIMSS frameworks influenced the mathematics and science components of this assessment, which expanded to include online formats and new subjects like science and foreign languages from 2021–2022 onward. (<i>Hungary</i>)	Israel’s national assessments of reading—originally known as Meitzav and more recently Tmunat Matzav—were developed and aligned in part with the PIRLS framework. (<i>Israel</i>)
Insights from TIMSS informed Ireland’s National Assessments of Mathematics and English Reading (NAMER), influencing both design and methodology. The frameworks and questionnaires from TIMSS were explicitly referenced in curriculum redevelopment and in shaping Ireland’s broader national strategy for literacy, numeracy, and digital literacy. (<i>Ireland</i>)	PIRLS frameworks and findings strongly influenced the design of Ireland’s National Assessment of English Reading, as well as standardized reading tests used in primary schools. The PIRLS-aligned structure contributed to literacy reforms and was cited in Ireland’s national literacy strategy. (<i>Ireland</i>)
TIMSS/PIRLS as role model for national assessments	
TIMSS (4)	PIRLS (5)
Armenia adopted the TIMSS methods and procedures to develop national assessments across a wide range of subjects: Armenian language/literature/history (HAAS), sciences (BAAS), and foreign languages (OLAS). (<i>Armenia</i>)	PIRLS was foundational in the design and institutionalization of the National Assessment of Basic Competencies (NABC). (<i>Hungary</i>)
Chile’s national standardized testing system (Simce) was influenced by TIMSS methodologies, including its evaluation framework, use of open-ended questions, and coding procedures. (<i>Chile</i>)	PIRLS findings were compared with the national assessment framework, contributing to adjustments in reading curriculum and assessments, especially around phonics instruction. (<i>England</i>)
TIMSS frameworks directly shaped Italy’s national assessment structures, particularly due to their strong curricular alignment. (<i>Italy</i>)	PIRLS helped shape debates on educational standards, contributed to the introduction of educational standards testing, and supported Austria’s evidence-based policymaking. (<i>Austria</i>)

Increased or informed funding	
TIMSS (3)	PIRLS (1)
TIMSS results, especially from the computer-based assessment (CBA), were explicitly used to justify increased investment in digital education infrastructure. (<i>Albania</i>)	The government allocated additional funding to primary schools for establishing class and central libraries and hiring teacher-librarians to support student reading and learning activities. (<i>Hong Kong SAR</i>)
TIMSS data contributed to national budget planning during the development of the 2014 core curriculum for basic education. (<i>Finland</i>)	n/a
TIMSS data have also been used in federal STEM grant policy and in setting funding priorities aimed at reducing achievement gaps, especially those related to gender, race, and socioeconomic status. (<i>USA</i>)	n/a
Monitor student progress over time or system monitoring	
TIMSS (21)	PIRLS (28)
Iran emphasized that TIMSS has replaced a lacking national system for monitoring education quality. It is used consistently to track achievement trends and has significantly influenced policymaking, curriculum reform, and the public discourse. (<i>Iran, Islamic Republic of</i>)	Germany integrates PIRLS as part of a national education monitoring strategy that includes both national and international assessments. PIRLS results are used at both school and policy levels. (<i>Germany</i>)
New Zealand explicitly described using TIMSS as a critical tool for long-term system monitoring in mathematics and science. (<i>New Zealand</i>)	Jordan has embedded PIRLS into a long-term system monitoring and reform strategy. The 2021 results serve as a baseline for upcoming reforms through 2031. (<i>Jordan</i>)
Chile clearly linked its use of TIMSS to monitoring changes in student learning over time and validating education policies. (<i>Chile</i>)	Poland views PIRLS as an essential component of education system monitoring due to the lack of other representative data for Grades 3 and 4. (<i>Poland</i>)
Comparison with other countries	
TIMSS (20)	PIRLS (18)
International comparisons—especially with neighboring EU countries—are cited by Austria as one of the two main purposes of participating in TIMSS. This comparative lens helped trigger debates on education quality and reforms and remains central to Austria's use of TIMSS for evaluating system performance. (<i>Austria</i>)	Australia highlighted PIRLS's value as a source of internationally comparative data on Year 4 reading literacy, used alongside NAPLAN to contextualize national student achievement. PIRLS results are viewed as important for evaluating Australian students' performance relative to peers globally. (<i>Australia</i>)
Hong Kong uses TIMSS as a benchmarking tool to assess student achievement against international standards. These comparisons are pivotal in shaping education policies, curriculum revisions, and public accountability. (<i>Hong Kong SAR</i>)	Singapore uses PIRLS to monitor student reading proficiency and benchmark student groups against international standards. These comparisons are central to developing targeted policy interventions and identifying areas needing improvement (e.g., reading for pleasure and online literacy skills). (<i>Singapore</i>)
England leverages TIMSS to identify high-performing and fast-improving education systems. Insights from these comparisons have directly influenced national reforms in mathematics and science curricula and informed major initiatives like the Maths Hubs program and teacher training materials developed by national education centers. (<i>England</i>)	Cyprus disseminates PIRLS results broadly, and these trigger public and stakeholder discussion centered on comparisons with other countries. The country uses findings from PIRLS, TIMSS, and PISA to develop policy recommendations for improving education outcomes. (<i>Cyprus</i>)

Within-country comparisons	
TIMSS (6)	PIRLS (2)
The U.S. extensively used TIMSS for state-level comparisons, allowing 18 states to participate as separate entities in one or more cycles. TIMSS data were also used to compare sub-groups by race/ethnicity, gender, and achievement levels, which helped identify widening gaps and inform both state-specific and national education reforms. (<i>USA</i>)	Canada used PIRLS data to compare student achievement both across provinces and within specific subgroups. Provincial Ministries use PIRLS to assess how their curricula are implemented, while researchers rely on the data to analyze subgroup performance—supporting policy and educational planning across diverse regional education systems. (<i>Canada</i>)
Italy emphasized territorial differences in educational outcomes, and TIMSS results informed the <i>RI.DI.TE. program</i> , aimed at reducing regional disparities and addressing dropout risks. The data are used to produce thematic reports for in-depth regional analysis and to guide place-based policy interventions. (<i>Italy</i>)	The U.S. used PIRLS to analyze achievement gaps between subgroups, such as by gender, highlighting disparities like girls outperforming boys. PIRLS findings are widely accessed and used by media, policymakers, and researchers to study demographic patterns in reading literacy across the country. (<i>United States</i>)
Abu Dhabi used TIMSS to produce school-specific reports for both public and private institutions, enabling comparison across curricula and school types. These reports guided instructional improvements and curriculum calibration in one of the world's most diverse school systems. (<i>Abu Dhabi</i>)	n/a
Use in teacher training and professionalization	
TIMSS (22)	PIRLS (26)
Bosnia & Herzegovina used TIMSS 2019 results extensively in teacher workshops, curriculum reforms, and the development of teaching strategies, especially in the West-Herzegovina Canton. Pedagogical institutes collaborated with educators to implement TIMSS-informed teaching methods, while the EU-funded Education for Employment project created professional development models based on TIMSS insights. (<i>Bosnia & Herzegovina</i>)	Following the PIRLS 2016 results, Belgium Flanders launched several targeted professional development programs like <i>Prioritaire nascholing</i> and <i>Lezen op school</i> . These programs focused on improving reading comprehension instruction, alongside teacher training initiatives informed by PIRLS follow-up studies and expert reviews. (<i>Belgium, Flemish</i>)
The Czech Republic developed a teacher-focused publication of TIMSS items with coding guides, instructional comments, and project ideas, along with organizing seminars and distributing resources to schools. These materials support methodological improvements and are directly applied in classrooms to address learning difficulties highlighted by TIMSS. (<i>Czech Republic</i>)	In Malta, PIRLS results directly informed professional development sessions in schools. Literacy specialists collaborated with teachers to improve comprehension instruction, and teacher education programs were updated to reflect PIRLS insights, especially on using more challenging texts. (<i>Malta</i>)
Kazakhstan used TIMSS data to reform teacher professional development programs, including initiatives by the <i>National Center for Professional Development "Orleu."</i> The TIMSS framework informed training modules to help teachers implement international assessments and align instructional practices with global standards. (<i>Kazakhstan</i>)	Jordan developed comprehensive teacher guides based on PIRLS item analysis, focusing on common student errors and instructional strategies. These guides are used in teacher education programs and form a key part of the country's broader education reform strategy. (<i>Jordan</i>)
Develop or improve teaching material	
TIMSS (9)	PIRLS (25)
The Y. Altynsarin <i>National Academy of Education</i> used the TIMSS conceptual model to develop educational programs and school textbooks. Additionally, the <i>National Center for Professional Development</i> created teaching support programs specifically aligned with international assessment expectations. (<i>Kazakhstan</i>)	Slovenia developed diagnostic materials for reading literacy and implemented large-scale training programs where teachers used PIRLS 2016 data to design school-specific studies and adjust instruction accordingly. (<i>Slovenia</i>)
TIMSS participation led to the development of items inspired by released TIMSS questions and remedial materials to support teachers and subject committees. These were used to enrich teaching content and directly address learning weaknesses identified through TIMSS data. (<i>Palestinian National Authority</i>)	Following the drop in PIRLS performance, Flanders initiated professional development programs, commissioned practical teaching guides, and launched a long-term Reading Offensive to improve reading instruction. (<i>Belgium, Flemish</i>)
Slovakia's national institute prepared mathematics and science textbooks based on released TIMSS items from 2007 and 2011. These were explicitly intended as inspirational teaching resources for schools to enhance classroom practice. (<i>Slovak Republic</i>)	PIRLS frameworks and released items were incorporated into primary language textbooks and teachers' editions, and professional development sessions were held to introduce these tools and support classroom use. (<i>Iran, Islamic Republic of</i>)

Increased or improved digital learning	
TIMSS (4)	PIRLS (2)
Qatar integrated educational software into curricula for problem-solving and investigation, aligned with TIMSS tasks, and moved decisively toward e-learning, training students across subjects to use digital devices and educational platforms in response to the digital nature of TIMSS and global trends. (<i>Qatar</i>)	Russia used PIRLS and ePIRLS to shift from oral to written literacy assessments, expand the curriculum's focus on informational texts, and prompted debate and reflection on how to combine traditional and digital reading strategies, with digital reading becoming an emerging area of pedagogical development. (<i>Russian Federation</i>)
Sharjah/UAE used TIMSS insights to introduce adaptive digital content in classrooms, refine digital instruction strategies, and align pedagogy with international best practices, with digital learning serving as a strategic focus for educational reform. (<i>Sharjah/UAE</i>)	n/a
Bahrain prepared students for TIMSS 2023's digital format by enhancing e-learning readiness, familiarizing students with digital assessments, and emphasizing online platforms—especially in response to the COVID-19 shift to remote learning. (<i>Bahrain</i>)	n/a
Introduced or increased STEM activities	
TIMSS (10)	PIRLS (n/a)
Montenegro identified low TIMSS results as a catalyst to organize STEM seminars and workshops for lower-primary teachers, using TIMSS to prioritize systemic professional development and activity-based learning in science and mathematics education. (<i>Montenegro</i>)	n/a
Lithuania responded to TIMSS insights by equipping all schools with science labs and launching 10 open-access STEAM centers that provide both formal and nonformal science and technology education, as well as establishing a <i>Science and Technology Popularization Centre</i> . (<i>Lithuania</i>)	n/a
England embedded TIMSS findings into the <i>National STEM Learning Centre's</i> professional development resources, supported the <i>Teaching for Mastery</i> program through its network of <i>Maths Hubs</i> , and aligned science curriculum revisions with TIMSS frameworks. (<i>England</i>)	n/a
Increased technical or scientific equipment	
TIMSS (3)	PIRLS (n/a)
Lithuania used TIMSS findings to equip all primary and basic schools (Grades 5–8) with science laboratories and also established 10 open-access STEAM centers, providing lab facilities for both formal and nonformal education. These infrastructure enhancements were directly motivated by insights from TIMSS 2011 onward. (<i>Lithuania</i>)	n/a
Albania committed to increasing funding for ICT infrastructure in primary education, in response to insights gained from TIMSS and its implementation as a computer-based assessment, aiming to expand digital learning tools and close the digital gap in mathematics and science education. (<i>Albania</i>)	n/a
Bahrain undertook efforts to familiarize students with digital tools and platforms, including the provision of digital activities and infrastructure for e-learning after participating in the TIMSS 2023 digital assessment, especially to enhance preparedness and equity across schools. (<i>Bahrain</i>)	n/a

Data source for researchers	
TIMSS (16)	PIRLS (20)
Researchers in Germany can access TIMSS data through a designated research data center (FDZ), and the data is used for academic theses and journal articles. Major national initiatives in science and mathematics education (e.g., SINUS, DZLM) are informed by TIMSS findings. (<i>Germany</i>)	PIRLS data in England has been widely used by researchers to explore the role of phonics instruction, examine international assessment biases, and study reading attainment. A specific project, <i>PIRLS for Teachers</i> , helped academics work directly with teachers to interpret PIRLS data for improving classroom pedagogy. (<i>England</i>)
TIMSS data is a widely used evidence base for researchers in South Africa and internationally. The data informs scholarly work—journal articles, books, theses—and is made accessible through the TIMSS SA website and national dissemination by the Department of Basic Education. (<i>South Africa</i>)	PIRLS is part of Germany's national education monitoring strategy and supports evidence-based regulation and instructional practices. Results are used in national education reports, and the strong research focus has led to the creation of professorships in empirical education research. (<i>Germany</i>)
TIMSS 2019 data has been used extensively by researchers across administrative units, including university-based studies, thematic workshops, and regional collaborations. APOSO published multiple secondary analyses (e.g., on preschool education and homework), encouraged systematic use of TIMSS for empirical research, and launched initiatives pairing experts to train teachers using TIMSS items as guides. (<i>Bosnia & Herzegovina</i>)	Slovenian researchers have incorporated PIRLS into regional conferences, national strategies, and multiple projects to train educators in reading literacy. Teachers have been taught to analyze PIRLS data using specialized tools (like RALSA), and design school-specific studies for instructional improvement. (<i>Slovenia</i>)
Increase in public awareness of education or students' performance	
TIMSS (9)	PIRLS (22)
Cyprus ensures broad dissemination of TIMSS results through national reports and presentations. These consistently stimulate extensive public debate in local media, particularly regarding how students compare internationally. (<i>Cyprus</i>)	Improved PIRLS 2016 results generated positive media coverage in Ireland, especially following concerns about student reading skills from earlier international assessments. PIRLS has played a role in shaping Ireland's literacy strategies and has been used to evaluate national education reforms, thereby contributing to public discourse on reading literacy. (<i>Ireland</i>)
TIMSS results generated substantial discourse among academics, politicians, and the general public, raising awareness about the need to improve mathematics and science education. Public concern over declining scores between 2003 and 2011 spurred national curriculum and assessment reforms. (<i>Malaysia</i>)	The significant decline in PIRLS 2016 scores received extensive and prolonged coverage in all Flemish newspapers and media, sparking national debates about the education system. This public concern led to follow-up research, media campaigns, and major reading-focused reforms, including a government-funded Reading Offensive continuing through 2030. (<i>Belgium, Flemish</i>)
TIMSS results in Hong Kong are widely used not only to inform policy but also to raise public awareness. Their release regularly prompts discussions among educators, parents, policymakers, and the public, fostering accountability and collaborative improvement in education. (<i>Hong Kong SAR</i>)	PIRLS data have been widely used in public debates in Austria, particularly around issues like students' reading deficits and the relationship between reading achievement and socioeconomic background. These discussions have informed reforms and the professional development of teachers, with the Ministry's Coordination Center for Reading playing a central role in disseminating results to the public and educators. (<i>Austria</i>)

- Note: Numbers in brackets indicate the number of education systems that mentioned the specific category for TIMSS (T) and PIRLS (P). Category "Introduced or evaluated reforms or special support programs" is grouped under two core areas because it is related to both.

APPENDIX B

The analysis followed an inductive qualitative content analysis approach, drawing on established coding procedures as outlined by Miles and Huberman (1994) and Saldaña (2015). All data extractions and codings were conducted manually using structured tables rather than specialized qualitative data analysis software.

The methodological approach consisted of five main stages:

Data extraction

The section on “impact and use of data” was extracted from each country’s encyclopedia chapter and compiled into a structured table, organized by education system. This step ensured that all relevant information was consolidated in a consistent format, facilitating comparison and thematic analysis. It avoids any pre-selection of information and keeps the entire relevant content as presented in each country’s encyclopedia chapter.

Qualitative content analysis: Defining core areas

In the absence of existing core areas, a strict deductive data-driven approach was adopted: All extracted text segments were systematically reviewed and coded to identify concrete instances of data use and observed or claimed impact. The analysis focused on identifying both unique, country-specific examples and cross-cutting themes that appeared in multiple national reports. The core areas were derived from recurring patterns in the texts and designed to accommodate both national variations and shared international experiences.

Development and application of the categorization framework

Building on the broad core areas, an inductive coding and categorization framework was developed to capture the most salient types of impact and use reported by countries in relation to TIMSS and PIRLS. The initial coding scheme was developed during a first coding pass of the TIMSS Impact and Use sections by Author 1. Texts of education systems were processed in alphabetical order. Whenever a new aspect of impact or use was identified, a new code was added to the existing list; in rare cases, existing codes were revised to better reflect the emerging material. Subsequent country texts were then examined to determine whether the existing codes were applicable, and new codes were added where necessary.

This process resulted in a coding matrix in which the rows represented countries and the columns represented the identified codes. A value of one was assigned where a code applied to a country-specific text, and a value of zero where it did

not. After the initial TIMSS coding scheme had been developed, the codes were applied to the PIRLS Impact and Use sections. The coding scheme proved to be highly stable, and no additional categories had to be developed for PIRLS. However, some TIMSS categories did not have an equivalent in the PIRLS material, such as the category related to STEM education.

In a subsequent step, the coding matrix was cleared and Author 2 independently coded the country-specific texts using the established coding scheme. Where Author 2 identified an additional aspect that was not yet captured, the corresponding code was incorporated into the coding scheme. The coding decisions of Author 1 and Author 2 were then compared, and discrepancies were discussed and resolved. This process resulted in a consolidated coding table.

The finalized framework enabled the systematic classification of a wide range of practices and outcomes associated with the use of TIMSS and PIRLS data. Based on the detailed codes, broader core areas were derived. The complete set of categories is presented in [Table 1](#) and [Figure 1](#).

Reliability and arbitration

To ensure the reliability and consistency of the analysis, the definition of categories and the application of these categories were conducted independently by both authors of this brief. Any discrepancies in the development of the category system or the assignment of country cases to categories were resolved through an arbitration process. Only if consensus was achieved was the category or its application accepted. This process involved in-depth discussions to refine category definitions and clarify ambiguous or inconsistent statements in the original texts. For TIMSS, the agreement between both authors was 97 percent, for PIRLS, the agreement resulted in 82.5 percent. This could be due to the fact that the category system was initially developed for TIMSS and then was applied to PIRLS, leading to some ambiguity in the PIRLS categorization.

Quality assurance

As an additional more experimental quality assurance measure, all country sections were reviewed using ChatGPT-4. This tool was employed to independently verify the classification of cases into broad core areas to ensure completeness and consistency in the application of the categorization framework. ChatGPT was not used to replace human judgement. As a result, ChatGPT generated categories that closely mirrored those developed independently by the authors. The application of these categories also showed no major deviations.