

IEA Research_ and Development_

Enhancing Questionnaire Relevance in Low- and Middle-Income Countries:

A Comparative Analysis of Regional Assessment Frameworks
and IEA Instruments in Primary Education

Oliver Neuschmidt¹
Clara Beyer¹
Sarah Howie²

¹International Association for the Evaluation of Educational Achievement

²Stellenbosch University, Stellenbosch, South Africa

DOI: https://doi.org/10.58150/rd_4_4

Summary

This IEA Research and Development project addresses a key challenge for international large-scale assessments (ILSAs): ensuring that context questionnaires remain internationally comparable while adequately reflecting the realities of low- and lower-middle-income countries (LMICs).

The project compared questionnaires and frameworks from IEA studies (TIMSS, PIRLS, and LaNA) and regional assessments, including ERCE, PASEC, SACMEQ, SEA-PLM, and AMPL, while also drawing on insights from PISA for Development (PISA-D). Based on this review, it identified relevant context indicators and developed a set of proposed questionnaire items across student, teacher, and school levels, together with a suitability classification to guide their future use and adaptation.

The findings show that many constructs can be measured effectively using existing IEA items, demonstrating the robustness and transferability of established IEA instruments. Student questionnaires required somewhat greater adaptation, and additional indicators were needed to better capture LMIC-specific conditions such as household circumstances, language environments, and health, as well as teacher labor conditions and educational pathways. The project recommends a modular questionnaire design consisting of a stable core of internationally comparable items complemented by context-sensitive extensions. These findings provide practical guidance for future IEA studies and other initiatives collecting comparable context information in LMICs, including for SDG 4 monitoring.

This work was supported in whole or part by the IEA Research and Development (R&D) Funds, which aim to advance and improve the science and methodology of IEA studies, ensuring that they remain at the forefront of international large-scale assessments in education. Funding is drawn from an overhead on participation fees across IEA studies. This does not constitute endorsement of the contents, which reflects the views only of the authors. IEA cannot be held responsible for any use which may be made of the information contained therein.

© Stichting IEA Secretariaat Nederland 2026. All commercial rights are reserved. A non-exclusive and non-commercial license was granted to the fund recipient to use the intellectual property for future purposes.

This work is licensed under a Creative Commons Attribution - NonCommercial 4.0 International License, which permits noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>.

Images or other third-party material in the work are included in the Creative Commons license unless indicated otherwise in a credit line to the material. If any material is not included in the work's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you must obtain permission directly from the copyright holder.

Enhancing Questionnaire Relevance in Low- and Middle-Income Countries:

A Comparative Analysis of Regional Assessment Frameworks and IEA Instruments in Primary Education

Oliver Neuschmidt, IEA Hamburg – oliver.neuschmidt@iea-hamburg.de

Clara Beyer, IEA Hamburg – clara.beyer@iea-hamburg.de

Sarah Howie, Stellenbosch University – sarahhowie@sun.ac.za

Acknowledgements

The authors would like to thank colleagues who contributed to the development of this report through technical discussions, review, information gathering, and editorial support.

In particular, we acknowledge:

- Katherine Reynolds for contributions to project discussions and strategic decisions
- Dana Kelly for detailed review and constructive comments
- Chen-Ho Lee for information collection, editing and formatting support
- Etienne Ouedrigo (LaNA NRC for Burkina Faso), Dr. Abdel Kareem Badran (LaNA NRC for Egypt), Julius Adedaja (LaNA NRC for Nigeria), Dr. Kanwal Ejaz (LaNA NRC for Pakistan), Dr. Mohammed Matar (LaNA NRC for the Palestinian Authority), Alioune Badara Diop (LaNA NRC for Senegal) and Massar Diop (LaNA Senegal) for their review and their valuable comments of the suggested questionnaire items sets compiled for this projects

Any remaining errors or omissions remain the responsibility of the authors.

Table of Contents

1	Executive Summary	11
2	Introduction	13
2.1	Background and Motivation.....	13
2.2	Overview of International and Regional Learning Assessments at the End of Primary Education considered for this Project.....	14
2.2.1	TIMSS and PIRLS (IEA)	14
2.2.2	LaNA (IEA)	15
2.2.3	AMPL (Assessment for Monitoring Learning)	15
2.2.4	ERCE (UNESCO OREALC)	15
2.2.5	PASEC (CONFEMEN)	15
2.2.6	SEA-PLM (Southeast Asia Primary Learning Metrics)	16
2.2.7	SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality)	16
2.2.8	PISA (OECD)	16
2.2.9	Relevance for the Present Project	16
3	Problem Setting, Aims and Research Questions	20
3.1	Problem Setting	20
3.2	Aims of the Study and Research Questions	20
4	Literature and Framework Review	23
4.1	Educational and Contextual Factors in LMICs.....	24
4.1.1	Student-Level Factors	24
4.1.2	Teacher and Classroom-Level Factors	28
4.1.3	School-Level Factors.....	31
4.1.4	SDG-Related Information Needs	34
4.2	Questionnaire Frameworks	37
4.2.1	Conceptual and Theoretical Frameworks in International Large-Scale Assessments.....	37
4.2.2	Historical Foundations: From OTL to System Models	37
4.2.3	IEA TIMSS & PIRLS Questionnaire Frameworks.....	38
4.2.4	Regional and LMIC-Focused Frameworks.....	41
4.3	From Frameworks to Questionnaires	42
5	Methodology	43
5.1	Data Sources and Materials Used.....	43
5.2	Framework Comparison and Analytical Approach.....	45
5.2.1	The Analyzed Frameworks	46
5.2.2	Level of Comparison (Domains, Constructs, Thematic Areas).....	47
5.2.3	Alignment of Terminology (Mapping, Harmonization, Clustering) and LMIC-Focused Frameworks	49
5.2.4	Identification of Common vs Complementary Themes	49
5.3	Conceptual Review, Indicator Construction, and Item Selection	49
5.3.1	Database Selection and Scoring Strategy	50
5.4	Item Identification, Selection, and Adaptation	53
5.5	Empirical Screening on Question and Indicator Level	55
5.5.1	Scope of Verification and Data Preparation	56
5.5.2	Handling of Achievement Outcomes and Data Linkage	57
5.5.3	Association Measure and Statistical Techniques	57
5.5.4	Screening Criteria, Reliability and Validity Considerations	58
5.6	Qualitative Input from LMIC Representatives	59
5.6.1	Design of the Review Instrument.....	59
5.6.2	General Questionnaire-Level Review	60
5.6.3	Implementation.....	60

5.7	Item Suitability Assessment	60
5.7.1	Psychometric Item Suitability	60
5.7.2	Expert Review-Based Item Suitability	62
5.7.3	Treatment of Newly Developed or Substantially Modified Items	62
5.7.4	Overall Item Suitability	63
5.8	SDG Indicator Evaluation	63
5.9	AI-Assisted Procedures	64
6	Results.....	65
6.1	Evidence from the Literature and Areas for Strengthening in IEA ILSA Frameworks	65
6.2	Communalities and Differences across Assessment Frameworks (RQ1)	68
6.2.1	Comparative Analysis of Regional Frameworks	68
6.2.2	Comparison of Regional Frameworks with IEA Frameworks	69
6.2.3	Similarities and Differences	76
6.3	Indicator Identification and Alignment Across Assessments (RQ2)	83
6.3.1	Summary of Indicator Mapping to the Analytical Framework	83
6.3.2	Overview of Retained Indicators	84
6.3.3	Alignment with SDG Indicators and Identified Gaps	95
6.4	Operationalization of Indicators	99
6.4.1	Student-Level Operationalization	100
6.4.2	Teacher-Level Operationalization	105
6.4.3	School-Level Operationalization	110
6.4.4	Cross-Level Patterns in Sources and Adaptation	115
6.5	Empirical Characteristics of Proposed Indicators and Selected Questions	116
6.6	Empirical Characteristics of Retained Questions	117
6.6.1	Overview of Empirical Screening Results	117
6.6.2	Association with Achievement	118
6.7	Qualitative Review of Retained Questions	119
6.7.1	Analytical Scope and Data Sources	119
6.7.2	Perceived Relevance and Acceptance of Retained Questions	119
6.7.3	Overall Evaluation of Suggested Questions and Indicators	120
6.7.4	Missing Elements as Reported by LaNA 2023 Linking Study NRCs	120
6.7.5	NRC Comments Regarding Individual Questions	121
6.7.6	Summary of the Qualitative Findings	122
6.8	Item Suitability Assessment	122
6.8.1	Overview of Item Suitability Findings	123
6.8.2	Item Suitability Patterns	127
7	Discussion and Implications.....	129
7.1	Synthesis of Findings and Answer to Research Questions	129
7.1.1	Common and Distinct Themes Across Assessment Frameworks (RQ1)	129
7.1.2	Identification of Robust and Policy-Relevant Indicators (RQ2)	130
7.1.3	Operationalization of Indicators with Existing Questionnaire Items (RQ3)	132
7.1.4	Item Evaluation and Suitability Review (RQ4)	134
7.1.5	Overall Synthesis	135
7.2	Methodological Reflections and Limitations	135
7.3	Implications for Questionnaire Design	136
7.3.1	Coverage and Alignment with Conceptual Frameworks	136
7.3.2	Extension to Emerging and Evolving Domains	137
7.3.3	Balancing Standardization and Contextual Adaptation	137
7.3.4	Questionnaire Design Across Respondent Groups	137
7.3.5	Use of Optional and Context-Specific Items	137
7.4	Policy and Analytical Use	138

7.4.1	Policy Relevance and Feasibility (NRC Perspective).....	138
7.4.2	Alignment with SDG Monitoring and Measurement Constraints.....	138
7.4.3	Interpretation of Indicators for Monitoring Purposes	138
7.5	Recommendations	139
7.5.1	Use and Validation of the Proposed Questionnaire Item Sets.....	139
7.5.2	Recommendations for Flagged Items	139
7.5.3	Recommendations for SDG-Oriented Measurement.....	143
7.5.4	Recommendations Based on Expert-Identified Indicator Gaps	143
7.5.5	Recommendations from Framework–Indicator Mapping	143
7.5.6	Practical Considerations for Questionnaire Design	144
7.6	Areas for Further Development and Outlook.....	144
7.6.1	Emerging Domains and Strengthening the Evidence Base.....	144
7.6.2	Administration Procedures, Timing, and Visual Design.....	144
8	Appendices	147
8.1	Appendix A1: Suggested Student Questionnaire Items.....	147
8.2	Appendix A2: Suggested Teacher Questionnaire Items	154
8.3	Appendix A3: Suggested School Questionnaire Items	161
8.4	Appendix A4: Full Indicator List.....	168
8.5	Appendix A5: Additional Indicators suggested by Country Experts	187
8.6	Appendix A6: Overview on Selected and Modifications.....	187
8.7	Appendix A7: AI Diligence Statement	228
8.7.1	Creation Diligence	228
8.7.2	Transparency Diligence	228
8.7.3	Deployment Diligence.....	229
	References	230

List of Tables

Table 2-1: Regional Assessments and Initiatives targeting end of primary education	18
Table 4-1: SDG indicators measurable via ILSA context questionnaires according to UIS (2025)35	
Table 5-1: Availability of Frameworks, Questionnaires, Data, and Documentation	44
Table 5-2: Source of Information Used for Thematic Analysis of Projects’ Contextual Framework/Data	47
Table 5-3: Analytical Grid Used for Analysis and Comparison of Frameworks	48
Table 5-4: Study-Specific LMIC Data used for Empirical Screening	55
Table 5-5: Number of Shared Countries Across Studies	56
Table 5-6: Summary of Screening Criteria and Thresholds for Background Variables	59
Table 6-1: LMIC-Relevant Contextual Indicators, Evidence support, and Coverage across LaNA and Regional Assessments.....	66
Table 6-2: System-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA studies	70
Table 6-3: School-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA studies	71
Table 6-4: Class-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA Studies	73
Table 6-5: Student-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA Studies.....	75
Table 6-6: Overview of the Domains and Topics Found Across Curricular Domains on System, School, Class and Student Levels in Selected International and Regional Studies	79
Table 6-7: Proposed Framework Based on Outcomes of Analyses	82

Table 6-8: Indicator-Matching with Project Framework	84
Table 6-9: Student Level Indicators	85
Table 6-10: Teacher Level Indicators.....	88
Table 6-11: School Level Indicators	91
Table 6-12: SDG Indicators Measurable via ILSA Context Questionnaires and Coverage in R&D Maintained Indicators	96
Table 6-13: SDG Indicators useful as Disaggregation measures or Metric Prerequisites	98
Table 6-14: Student Level Mapping.....	101
Table 6-15: Teacher Level Mapping.....	105
Table 6-16: School Level Mapping	110
Table 6-17: Summary of Empirical Screening Results by Questionnaire Level.....	117
Table 6-18: Summary of NRC Acceptance Ratings by Questionnaire Level	119
Table 6-19: Overall Evaluation of Suggested Questions and Indicators (Average Ratings)	120
Table 6-20: Item Suitability Categorization	123
Table 6-21: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (Student Level).....	124
Table 6-22: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (Teacher Level).....	125
Table 6-23: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (School Level)	126
Table 6-24: Distribution of Item Suitability by Questionnaire Level.....	127
Table 7-1: Recommendation on Question Level (Students)	140
Table 7-2: Recommendation on Question Level (Teachers and Schools).....	141
Table 8-1: Full Indicator List, Student Level.....	168
Table 8-2: Full Indicator List, Teacher Level.....	173
Table 8-3: Full Indicator List, School Level.....	179
Table 8-4: Additional Indicators suggested by Country Experts	187
Table 8-5: Student Questionnaire Selected Items and Modifications	188
Table 8-6: Teacher Questionnaire Selected Items and Modifications.....	200
Table 8-7: School Questionnaire Selected Items and Modifications.....	215

List of Figures

Figure 2-1: Participating Countries by Assessment and World Bank Income Group	18
Figure 4-1: Model of an IEA Research Study (IEA 1998, p. 53; Travers and Westbury, 1989 in Plomp et al., 2003).	39
Figure 7-1: Visualization example from ERCE Grade 3 Questionnaire	145
Figure 7-2: Visualization example from TIMSS Grade 4 Questionnaire.....	145

Table of Abbreviations

Abbreviation	Full Form	Note
AMPL	Assessment for Monitoring Learning	Modular LMIC assessment for SDG 4.1.1 monitoring
CIPP	Context, Input, Process, Product	Evaluation model underlying the ERCE framework
CONFEMEN	Conférence des Ministres de l'Éducation des États et gouvernements de la Francophonie	Governing body of PASEC
EAS	Emphasis on Academic Success	Indicator cluster label; school climate construct
ECE	Early Childhood Education	Pre-primary participation; linked to SDG 4.2
EER	Educational Effectiveness Research	Academic field; multilevel causal models of school quality
ERCE	Estudio Regional Comparativo y Explicativo	Regional Assessment in Latin America
ESD	Education for Sustainable Development	Domain in ERCE framework
FIMS	First International Mathematics Study	Early IEA study (1959–1967)
FISS	First International Science Study	Early IEA study (1966–1973)
GAML	Global Alliance to Monitor Learning	UNESCO-led framework for SDG 4.1.1 measurement
GCED	Global Citizenship Education	Domain in ERCE and SEA-PLM frameworks
HIC	High-Income Country	World Bank income classification
HIV/AIDS	Human Immunodeficiency Virus / Acquired Immune Deficiency Syndrome	Health topic in SACMEQ/PASEC questionnaires
ICILS	International Computer and Information Literacy Study	IEA study on digital literacy
ICT	Information and Communication Technologies	Digital tools in education
IEA	International Association for the Evaluation of Educational Achievement	Governing body; conducts TIMSS, PIRLS, LaNA
ILO	International Labour Office	UN agency; cited on child labor statistics
ILSA	International Large-Scale Assessment	Generic term for comparative assessments like TIMSS, PIRLS
ISCED	International Standard Classification of Education	UNESCO education level classification
LaNA	Literacy and Numeracy Assessment	IEA assessment designed for LMIC contexts; end of primary
LLECE	Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación	UNESCO regional lab; Coordinates ERCE study
LMIC	Low- and Lower-Middle-Income Country	World Bank income classification; primary target context
MAIL	Media and AI Literacy	Forthcoming PISA 2029 framework domain
MILO	Monitoring Individual Learning Outcomes	AMPL predecessor study
MOD	Modification level (1–3)	Classification of item adaptation extent
NRC	National Research Coordinator	Country-level ILSA study manager
OTL	Opportunity to Learn	Curriculum coverage construct; foundational to IEA
PASEC	Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN	Regional assessment in Francophone Sub-Saharan Africa
PIAAC	Programme for the International Assessment of Adult Competencies	OECD adult competency study
PILNA	Pacific Islands Literacy and Numeracy Assessment	Regional assessment (excluded from this study)
PIRLS	Progress in International Reading Literacy Study	IEA Reading Comprehension Assessment; Grade 4 reading; every five years
PISA	Programme for International Student Assessment	OECD Assessment; 15-year-olds

Abbreviation	Full Form	Note
PISA-D	PISA for Development	OECD adaptation of PISA for LMIC contexts
PRIDI	Programa Regional de Indicadores de Desarrollo Infantil	IDB regional early childhood indicator program
QUIRC	Questionnaire Item Review Committee	PIRLS 2026 / TIMSS 2027 internal item review body
R&D	Research and Development	Project classification within IEA
RQ	Research Question	Numbered RQ1–RQ4 structuring the study
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality	Regional assessment in Southern & Eastern Africa
SDG	Sustainable Development Goal	UN 2030 Agenda; SDG 4 covers quality education
SEA-PLM	Southeast Asia Primary Learning Metrics	Regional assessment in Southeast Asia
SES	Socioeconomic Status	Composite measure of household/social position; key predictor
SIMS	Second International Mathematics Study	IEA study (1976–1987); foundational for OTL
SISS	Second International Science Study	IEA study (1980–1989)
SOS	Safe and Orderly School	Indicator cluster label; school climate construct
TALIS	Teaching and Learning International Survey	OECD teacher survey
TERCE	Tercer Estudio Regional Comparativo y Explicativo	Earlier cycle of ERCE
TIMSS	Trends in International Mathematics and Science Study	IEA Mathematics and Science assessment; Grades 4 & 8; every four years
UIS	UNESCO Institute for Statistics	UN body producing SDG education statistics
UMIC	Upper-Middle-Income Country	World Bank income classification
UNICEF	United Nations Children’s Fund	UN agency; cited on child well-being data
WB	World Bank	Partner organization; provided indicator priorities

Technical Terms

Term	Definition
CIPP model¹	Context–Input–Process–Product evaluation model; conceptual basis of the ERCE framework, structuring analysis around what conditions, resources, processes, and outcomes characterize a school system.
Curricular antecedents	Background conditions preceding curriculum delivery: socioeconomic status, school resources, system-level features. One of three main dimensions in the analytical grid used for framework comparison.
Curricular contexts	Immediate process-level conditions surrounding curriculum delivery: classroom climate, teaching practices, school organization.
Differential item functioning (DIF)²	A psychometric property where an item behaves differently for subgroups (e.g., gender, country) after controlling for the underlying construct, threatening cross-group comparability.
Double-shift schooling	Organizational arrangement in which two student groups use the same building in separate time slots, reducing effective instructional time. Common in LMIC contexts.

¹ Stufflebeam (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31–62). Kluwer.

² Holland and Wainer (2012). *Differential item functioning*. Lawrence Erlbaum.

Term	Definition
Dynamic model of educational effectiveness³	Model by Creemers & Kyriakides (2008) conceptualizing school and classroom processes as evolving and reciprocal, integrating multiple levels into an explanatory framework for learning outcomes.
Educational Prosperity model⁴	Conceptual framework of PISA-D; a life-course approach identifying prosperity outcomes (academic performance, health, well-being, attainment) and foundations for success across developmental stages.
Expert override	Mechanism whereby the research team retains an indicator that did not meet the quantitative scoring threshold, based on professional judgment about its LMIC importance.
Grade retention / over-age for grade	Requirement to repeat a school year rather than advance; linked to SDG 4.1.5. Proxied via age and grade repetition data.
Heyneman-Loxley effect⁵	Empirical finding (Heyneman & Loxley, 1983) that in lower-income countries, school resources explain more variance in achievement relative to family background than in wealthier countries. Remains debated.
Input-process-output model⁶	Early IEA systems model linking educational inputs through teaching processes to achievement outputs; predecessor to multilevel models.
Intended / Implemented / Attained curriculum⁷	Three-tier IEA curriculum model: national policy (intended) → what teachers actually teach (implemented) → what students have learned (attained).
Item pool / candidate items	Full set of questionnaire items compiled from multiple source assessments and reviewed before final selection and adaptation.
Minimum proficiency level⁸	Achievement threshold for SDG 4.1.1; defines whether a student has attained foundational competencies in reading and mathematics. AMPL is explicitly designed to measure this threshold.
Modification level (MOD 1–3)	Classification of item adaptation: MOD 1 = minimal change; MOD 2 = moderate adaptation; MOD 3 = substantial rewrite or new item. MOD 3 items cannot reuse original psychometric parameters.
Multi-grade teaching	One teacher simultaneously teaching students from multiple grade levels. Common in rural LMIC schools; flagged as a key contextual factor.
Multilevel framework / multilevel modeling⁹	Statistical and conceptual approach accounting for nested data structure (students within classrooms within schools within systems). Standard in ILSA analysis.
Omission rate / missingness	Proportion of respondents without a valid answer to an item. Values ≥20% are treated as a risk indicator suggesting comprehension difficulties or implementation problems.
Opportunity to Learn (OTL)¹⁰	The extent to which students have been exposed to curriculum content required to answer assessment items. A foundational IEA construct for valid cross-country comparison.
Plausible values¹¹	Multiple imputed achievement score estimates used in ILSA analyses to account for measurement uncertainty from partial test designs.

³ Creemers and Kyriakides (2008). The dynamics of educational effectiveness. Routledge.

⁴ Willms (2018). Learning divides: Using data to inform educational policy. UNESCO Institute for Statistics.

⁵ Heyneman and Loxley (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. *American Journal of Sociology*, 88(6), 1162–1194.

⁶ Scheerens (1990). School effectiveness and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1(1), 61–80.

⁷ Travers, K. J., & Westbury, I. (1989). The IEA study of mathematics I: Analysis of mathematics curricula. Pergamon Press.

⁸ Australian Council for Educational Research (2022). Minimum proficiency levels: Described, unpacked and illustrated. UIS.

⁹ Raudenbush and Bryk (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Sage.

¹⁰ McDonnell (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, 17(3), 305–322.

¹¹ Mislevy (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56(2), 177–196.

Term	Definition
Proxy measure / Wealth index¹²	A variable used to indirectly represent a hard-to-measure concept. Household possessions and housing materials serve as socioeconomic proxies in LMIC contexts; a wealth index aggregates these into a composite score.
Psychometric item suitability	Classification of whether an item shows acceptable empirical properties (low omission rate, sufficient variability) based on data from existing assessments. Distinct from expert-based evaluation.
Pupil–teacher ratio	Number of students per teacher. Relevant to SDG 4.c.2 and 4.c.4. Cannot be derived from questionnaires alone; requires enrollment and staff count data.
Rasch model¹³	A psychometric IRT model placing student ability and item difficulty on a common scale. Referenced through the LaNA 2023 Linking Study.
Respondent burden	Cognitive and time demands placed on a survey respondent. Minimized by simpler items and shorter scales, especially for primary-school students.
R² (coefficient of determination)	Proportion of variance in achievement explained by a background variable in a regression model. Values below 0.01 are flagged as negligible in the psychometric screening.
SDG 4¹⁴	Sustainable Development Goal 4: 'Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.' The overarching global education policy framework referenced throughout.
Scale / composite indicator	An aggregate score combining multiple items to measure a latent construct more reliably than a single item. Referenced for socioeconomic background indices and attitudinal scales.
Skip logic / routing	Conditional instructions directing respondents to skip questions based on prior answers. Avoided in student instruments due to higher implementation-error risk in LMIC settings.

¹² Smits and Steendijk (2015). The international Wealth Index (IWI). Soc Indic Res 122(1), 65-85.

¹³ Rasch (1960). Probabilistic models for some intelligence and attainment tests. Danish Institute for Educational Research.

¹⁴ United Nations (2015). Transforming our world: The 2030 agenda for sustainable development (A/RES/70/1). United Nations.

1 Executive Summary

This IEA Research and Development project addresses a central methodological and practical challenge in international comparative education: the development of context questionnaires that are both internationally comparable and locally meaningful, particularly in low- and lower-middle-income countries (LMICs). While international large-scale assessments (ILSAs) such as TIMSS and PIRLS provide well-established conceptual frameworks and instruments, their context instruments are primarily designed for higher-income education systems and may not comprehensively capture structural inequalities, diverse household conditions, and context-specific learning environments in LMIC contexts. The present study responds to this challenge by developing a systematic framework for the identification, adaptation, and evaluation of student-, teacher-, and school questionnaire items for use in LMIC contexts, drawing on regional assessments and LMIC-focused initiatives. The project addresses three overarching research questions: how existing international and regional questionnaire frameworks compare in their coverage of relevant constructs on framework and questionnaire level; how suitable indicators can be selected and operationalized for LMIC contexts; and how the resulting questionnaire items can be evaluated to inform their practical use.

The project builds on an extensive review and integration of existing end-of-primary questionnaire instruments from both international and regional assessment programs, including TIMSS, PIRLS, ERCE, PASEC, SACMEQ, SEA-PLM, and AMPL; in addition, selected constructs and items are informed by PISA-D (secondary level) where relevant. These sources provide a broad empirical and conceptual basis for constructing a comprehensive pool of candidate indicators across student, teacher, and school levels. The project draws on two complementary source layers: IEA frameworks and instruments (TIMSS, PIRLS, LaNA), which provide a common comparative reference structure and internationally established constructs, and regional and LMIC-focused assessments (ERCE, PASEC, SEA-PLM, SACMEQ, AMPL, PISA-D), which provide context-sensitive constructs and operationalizations grounded in LMIC educational contexts. The IEA frameworks serve as the initial organizing and alignment anchor for the questionnaire review, while regional and LMIC-focused assessments are systematically incorporated to strengthen contextual relevance and applicability in LMIC settings.

Methodologically, the study follows a structured multi-step approach. First, indicators are mapped across frameworks to identify common constructs and gaps in coverage, with particular attention to domains that are underrepresented in international instruments but highly relevant in LMIC contexts. Second, selected indicators are matched with corresponding questions across all assessment sources and the most suitable items are identified based on predefined criteria (e.g., complexity, respondent burden, and contextual relevance); these items are then systematically adapted to improve clarity, cultural appropriateness, and suitability for primary-grade respondents.

Third, the adapted items are evaluated through a comprehensive item suitability framework that integrates multiple sources of evidence. This includes expert assessments from National Research Coordinators, empirical analyses of item characteristics such as differentiation (degree to which the item was assumed to distinguish between respondents) and missing rates, and an evaluation of policy relevance (degree to which the information was relevant for policy or system monitoring). By combining qualitative and quantitative evidence, the study develops a classification scheme that identifies items as robust, adaptable, or requiring further revision.

These classifications are subsequently synthesized into an overall guidance rating intended to support future questionnaire development and adaptation processes.

The findings reveal distinct patterns across educational levels. At the school level, most indicators can be directly sourced from IEA instruments with minimal adaptation, reflecting the relative higher stability and comparability of institutional and structural characteristics across contexts. Teacher-level indicators show more adaptation requirements, particularly in areas related to training pathways, employment conditions, and professional development, which vary considerably across countries. In contrast, student-level indicators require the most substantial adaptation. This is driven by the need to align item wording with the cognitive and linguistic capacities of younger respondents and to incorporate context-relevant measures of home environments and socioeconomic conditions.

The study also presents important conceptual considerations regarding the measurement of socioeconomic status. While traditional indicators such as parental education remain central, LMIC contexts often rely on proxy measures such as housing conditions, household composition, and student engagement in work activities. Their inclusion is essential for capturing structural inequalities in resource-constrained contexts.

Overall, the project contributes a transparent and evidence-based methodology for questionnaire development in LMIC settings. By bridging international and regional assessment frameworks and systematically integrating expert and empirical evidence, it provides practical guidance for designing context questionnaires that are both policy-relevant and methodologically sound. The project has the potential to strengthen the capacity of LMIC-specific assessments such as LaNA to generate meaningful context data on learning conditions and inequalities, while maintaining the level of comparability required for cross-national analysis and global monitoring efforts, including those related to context factors under Sustainable Development Goal 4 (SDG 4).

Based on these findings, several practical recommendations emerge for international comparative assessments operating in LMIC contexts: Indicator selection should draw on both internationally established constructs and regionally validated proxies, with source priority determined by construct validity and contextual fit rather than assessment origin. This is particularly important for socioeconomic status and related context domains (e.g., basic resources, health, and formal teacher education), where ensuring contextual validity is as important as maintaining cross-national comparability. Beyond this common LMIC core, additional domains may need to be considered to more comprehensively capture learning conditions, including process variables at the classroom level and emerging areas such as the use of digital tools and AI in teaching and learning. Moreover, questionnaire development for LMIC contexts should prioritize an adapted item design, particularly at the student level, including simplified language, the use of visuals, and adjusted administration procedures (e.g., reading items aloud where appropriate), as well as context-relevant response formats.

2 Introduction

2.1 Background and Motivation

IEA has been conducting internationally comparative studies of educational achievement for over 60 years, becoming a global gold standard for quality and reliable International Large-Scale Assessment (ILSA) data. Achievement results, however, are difficult to interpret without associated contextual data. The context questionnaires developed and administered by IEA gather crucial data on students' socioeconomic status (SES), learning environments, and educational resources, which are vital for understanding the broader context of students' academic performance.

These questionnaires, however, administered via IEA's well-established ILSAs such as TIMSS and PIRLS, are primarily designed—and function well—for the high and upper-middle income countries (HICs /UMICs) that have historically participated. IEA's mandate aims to facilitate improving education systems worldwide. There is now an increasing number of low and lower-middle income countries (LMICs) participating in IEA assessments. This study aims to address the significantly differing learning conditions in LMICs from those in HICs by suggesting enhancements and modifications to the context questionnaires administered via IEA ILSAs, thereby providing increased contextual relevance and validity for these countries and expanding IEA ILSA applicability and relevance worldwide.

LMICs have also increased their participation in regional assessments. Initiatives such as the CONFEMEN Programme for the Analysis of Education Systems (PASEC), which assesses learning achievement in Africa, particularly in Francophone sub-Saharan countries; UNESCO's Regional Comparative and Explanatory Study (ERCE), which assesses learning achievement in Latin American countries; the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), which carries out large-scale cross-national research studies in member countries in the Southern and Eastern Africa region; and the Southeast Asia Primary Learning Metrics (SEA-PLM), a regional learning assessment and capacity building program designed by and for Southeast Asian countries; have now contributed a rich knowledge base of contextual questionnaires and associated outcomes that are more targeted to the educational and social contexts in LMICs.

While this expanding body of regional evidence underscores the importance of contextualized measures, it also reflects the limitations of relying on constructs that were originally developed for, and validated in, high-income settings. More relevant and valid contextual questionnaires will strengthen IEA's capacity to capture the realities of education in LMICs, improve the explanatory power of background variables for student achievement, and enhance the utility of resulting data for national policymakers. Without such adaptation, essential factors shaping education in LMICs—such as access to basic infrastructure, household material conditions, teacher preparation, health and sanitation, absenteeism, and community-level influences—remain insufficiently measured.

As part of this work, particular consideration is given to the Sustainable Development Goals (SDGs), ensuring that proposed questionnaire item sets support LMICs in monitoring progress toward key education-related targets, including, for example, access to sanitation and health factors, equitable learning opportunities, and safe, well-resourced school environments.

The rationale for this study, therefore, was to systematically review the available materials from regional assessments, international initiatives, and scholarly literature to identify key constructs relevant to LMIC contexts and evaluate their alignment with existing IEA context instruments. This project seeks to lay an evidence-based foundation for enhancing IEA's context questionnaires and ensuring they better reflect the diverse realities of participating education systems.

2.2 Overview of International and Regional Learning Assessments at the End of Primary Education considered for this Project

This section provides an overview of the international and regional large-scale assessments that inform this study. While differing in scope, governance, and regional focus, these studies share a common objective: they assess student learning in the core foundational domains of mathematics and reading, in some cases alongside additional subjects such as science. Importantly, they also collect contextual information at home, class, and school levels through structured background questionnaires, which form the basis for the indicator development in this study. Except for PISA-D, all included studies and initiatives feature at least one target grade at the end of primary education.

Please note that while the Pacific Islands Literacy and Numeracy Assessment (PILNA) represents an important regional initiative at the end of primary education, it was not included in the comparative analysis due to its focus on a small number of geographically remote island systems and the resulting limited representativeness for the broader range of LMIC contexts considered in this study.

2.2.1 TIMSS and PIRLS (IEA)

The Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), conducted by the International Association for the Evaluation of Educational Achievement (IEA), serve as global benchmark assessments at the end of primary education.

TIMSS, conducted every four years, assesses mathematics and science in Grade 4 and Grade 8. It is grounded in a curriculum model distinguishing the intended, implemented, and attained curriculum, linking national curricular goals, classroom practices, and student achievement. The assessment framework combines subject-specific content domains with cognitive domains (knowing, applying, reasoning).

PIRLS, conducted every five years, assesses reading literacy at Grade 4. Its framework distinguishes between purposes for reading (literary and informational) and comprehension processes (Focus on and retrieve explicitly stated information, make straightforward inferences, Interpret and integrate ideas and information, and evaluate and critique content and textual elements).

Both studies include background questionnaires for students, teachers, and school principals. They also include a home/parent questionnaire, which, through the TIMSS 2023 cycle, was administered only at grade 4. These instruments collect contextual information on, for example, instructional practices, teacher preparation, school resources, leadership, student attitudes, early learning experiences, and home learning environments.

2.2.2 LaNA (IEA)

The Literacy and Numeracy Assessment (LaNA) was developed by the IEA and the TIMSS & PIRLS International Study Center as an assessment option tailored to countries that need timely, reliable, nationally representative, and internationally comparable data and contexts where existing international assessments may not adequately reflect lower proficiency ranges.

LaNA assesses emerging literacy and numeracy at the end of primary education. It does not operate under a separate standalone assessment framework; rather, its conceptual foundations are rooted in the TIMSS and PIRLS frameworks. The LaNA 2023 Linking Study established a psychometric link to the TIMSS/PIRLS scales, enabling international comparability while maintaining lower cognitive demand.

LaNA includes student and school questionnaires aligned with IEA contextual frameworks and adapted to LMIC contexts to a limited extent with focus on school conditions, instructional practices, student learning resources, and system-level factors relevant to foundational skill development.

2.2.3 AMPL (Assessment for Monitoring Learning)

AMPL is a modular assessment initiative designed to support monitoring of SDG indicator 4.1.1 on minimum proficiency levels in reading and mathematics. It focuses on foundational learning outcomes at key stages of primary education.

The conceptual orientation of AMPL is explicitly aligned with SDG measurement frameworks. Although narrower in scope than comprehensive ILSAs, AMPL includes contextual questionnaires at the student and school levels, capturing information on access to learning materials, instructional time, and basic school conditions relevant to minimum proficiency attainment.

2.2.4 ERCE (UNESCO OREALC)

The Estudio Regional Comparativo y Explicativo (ERCE), coordinated by UNESCO's Regional Bureau for Education in Latin America and the Caribbean, assesses student achievement in Grades 3 and 6 in reading, mathematics, and science.

ERCE is explicitly curriculum-based, aligning assessment content with national standards of participating countries. Beyond cognitive outcomes, ERCE places strong emphasis on socioemotional skills, Global Citizenship Education (GCED), and Education for Sustainable Development (ESD).

Background questionnaires are administered at the student, parent, teacher, and school levels. These instruments collect information on school climate, instructional practices, socioemotional development, equity-related factors, and home background characteristics.

2.2.5 PASEC (CONFEMEN)

The Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN (PASEC) operates in Francophone Sub-Saharan Africa and assesses competencies in the language of instruction and mathematics at the beginning and end of primary education.

PASEC applies a structured framework aligned with national curricula and emphasizes system diagnostics. A distinctive feature of the 2019 cycle was the inclusion of teacher assessments measuring subject knowledge and pedagogical skills.

Background questionnaires are administered to students, teachers, and school leaders, capturing information on teacher preparation, instructional time, classroom practices, school infrastructure, textbook availability, and governance conditions.

2.2.6 SEA-PLM (Southeast Asia Primary Learning Metrics)

SEA-PLM is a regional initiative led by Southeast Asian education ministries. It assesses Grade 5 students in reading, writing, and mathematics, with a distinctive writing literacy component.

The framework combines cognitive domains with a strong focus on regional priorities, including Global Citizenship. Background questionnaires at the student, teacher, school, and home levels gather information on instructional strategies, learning materials, school climate, parental involvement, and socioeconomic background.

2.2.7 SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality)

SACMEQ IV assessed Grade 6 students in reading and mathematics across participating ministries in Southern and Eastern Africa.

The study emphasizes system-level monitoring and equity analysis. Background questionnaires at the student, teacher, and school levels collect detailed information on school resources, textbook availability, teacher characteristics, classroom practices, and home conditions.

2.2.8 PISA (OECD)

PISA for Development (PISA-D) was developed by the OECD to extend the Programme for International Student Assessment (PISA) to low- and middle-income country (LMIC) contexts (OECD, 2018). The initiative aimed to enhance the relevance and validity of PISA for countries with substantially different educational and socioeconomic conditions than the higher-income countries that historically dominated participation in OECD assessments. PISA-D assessed 15-year-old students in reading, mathematics, and science, while adapting cognitive instruments to better capture lower proficiency levels and including an out-of-school assessment component (OECD, 2018).

Beyond the cognitive assessment adaptations, PISA-D is particularly relevant for the current research project because it represents one of the first major international large-scale assessment initiatives to systematically reconsider contextual questionnaire design for LMIC settings. The initiative acknowledged that background measures and contextual constructs originally developed primarily for higher-income countries may not fully capture the educational realities of LMICs (Willms, J., & Tramonte, L., 2015).

Although focused on secondary-age students, PISA-D provides potentially relevant contextual constructs at the student, teacher, school, and system levels that may inform the current research project. Background questionnaires cover areas such as school governance, instructional quality, teacher qualifications, learning time, student socioeconomic background, school resources, and system-level characteristics. The initiative is particularly relevant in relation to the adaptation and extension of socioeconomic and contextual measures for LMIC settings, which align closely with the objectives of the present study.

2.2.9 Relevance for the Present Project

The adaptation of contextual frameworks and questionnaire designs for diverse educational and socioeconomic settings has increasingly been discussed within the field of international large-

scale assessments (Rutkowski & Rutkowski, 2010; Sandoval-Hernández, Rutkowski, Tyler, & Daniel, 2019; Willms, J., & Tramate, L., 2015). These discussions emphasize that constructs and indicators originally developed primarily for higher-income countries may not adequately capture relevant dimensions of educational opportunity, home background, schooling conditions, and educational disadvantage in LMIC contexts. Consequently, several authors have argued for more context-sensitive approaches to questionnaire development, including the incorporation of regionally relevant indicators, alternative socioeconomic measures, and constructs that better reflect the educational realities of low-resource settings.

Across these assessments, contextual questionnaires are administered at multiple analytical levels—student, teacher, school, and in several cases home or parent level. While frameworks differ in structure and emphasis, they collectively cover domains such as socioeconomic background, instructional quality, school resources, governance, learning time, equity, and socioemotional development.

This diversity of conceptual frameworks and questionnaire designs provides the empirical basis for the indicator comparison and harmonization undertaken in this study. The objective is to identify indicators and corresponding questionnaire items that are relevant and applicable across all or most of these assessment contexts

Table 2-1 provides an overview of the main characteristics of the different assessments included in the current research project.

Table 2-1: Regional Assessments and Initiatives targeting end of primary education

Study / Initiative	Study Cycle (used)	Target Population / Grade	Subjects Assessed	Scope	Number of Participating Countries
TIMSS	2019	Grade 4	Mathematics; Science	Global	~58
PIRLS	2021	Grade 4	Reading literacy	Global	~57
LaNA	2023 Linking Study	End of primary education (Grade 4 to 6)	Reading; Mathematics	LMIC-focused, multi-regional	6 (in2023)
ERCE	2019	Grades 6	Reading; Mathematics; Science	Latin America and the Caribbean	16
PASEC	2019	Grade 5 or 6	Reading; Mathematics	Francophone Sub-Saharan Africa	14
SEA-PLM	2019	Grade 5	Reading; Mathematics; Global Citizenship	Southeast Asia	6
SACMEQ IV	2013–2014	Grade 6	Reading; Mathematics	Southern and Eastern Africa	14
AMPL ¹⁵	2023 administrations	End of primary (varies)	Reading; Mathematics (modular)	LMIC-focused	4 (+1 pilot)
PISA-D ¹⁶	2018	15-year-old students	Reading; Mathematics; Science	Selected LMICs (multi-regional)	7

Together, these studies represent complementary assessment approaches, ranging from globally standardized frameworks to regionally adapted and LMIC-focused instruments, providing the empirical basis for the integrated questionnaire development approach adopted in this study.

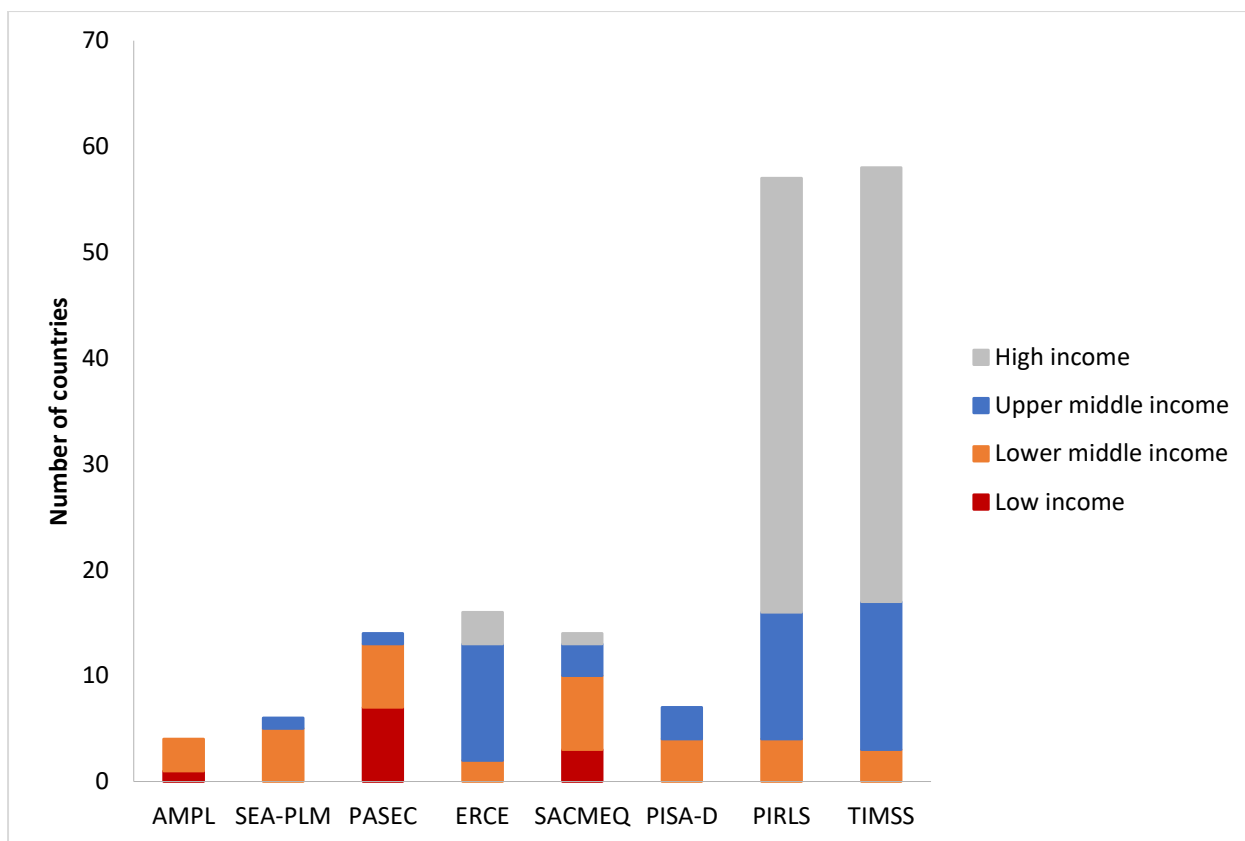
Figure 2-1 adds additional context by showing the number of participating countries by assessment and by World Bank income group.¹⁷ This provides insight into which segments of the global education landscape are primarily represented in different assessments and helps to identify the range of contexts that need to be targeted when developing questionnaire frameworks for LMIC settings.

Figure 2-1: Participating Countries by Assessment and World Bank Income Group

¹⁵ Based on predecessor study MILO (Monitoring Impacts on Learning Outcomes).

¹⁶ PISA-D assesses 15-year-olds (lower secondary); included here for its LMIC relevance.

¹⁷ World Bank income groups classify countries into low-, lower-middle-, upper-middle-, and high-income categories based on gross national income (GNI) per capita using the Atlas method. These classifications are updated annually and are intended for analytical and statistical purposes (World Bank (2025); see also <https://blogs.worldbank.org/en/opendata/world-bank-country-classifications-by-income-level-for-2024-2025>).



TIMSS and PIRLS are predominantly implemented in higher- and upper middle-income countries, whereas ERCE mainly concentrates on upper middle-income contexts. The remaining regional and LMIC-focused assessments are largely situated in lower middle- and low-income countries, with PASEC and SACMEQ showing the strongest representation of low-income systems. This distribution shows that assessments cover different segments of the income spectrum, underscoring the importance of integrating multiple sources when developing questionnaire frameworks applicable across diverse LMIC contexts. While TIMSS and PIRLS are proportionally concentrated in higher- and upper-middle-income countries, they still cover an absolute number of lower-middle-income countries comparable to several of the other assessments.

3 Problem Setting, Aims and Research Questions

3.1 Problem Setting

International and regional learning assessments increasingly rely on context questionnaires to support the interpretation of achievement results and to inform education policy and system monitoring. While International Large-Scale Assessments (ILSAs) such as TIMSS, PIRLS, and LaNA provide well-established questionnaire frameworks and instruments, concerns persist regarding the extent to which these instruments adequately capture the educational realities of LMICs. Contextual factors that are more strongly linked to achievement in LMIC settings may be underrepresented, insufficiently differentiated, or operationalized in ways that limit their usefulness for national policy dialogue and system-level analysis.

A growing number of regional assessment initiatives and globally oriented tools have emerged with an explicit focus on LMIC contexts. These initiatives often include context frameworks and questionnaire content that reflect region-specific priorities, implementation constraints, and policy needs. These frameworks are rarely systematically compared across initiatives, however, and their relationship to established international assessment frameworks remains only partially articulated. This represents an important opportunity to identify common constructs, complementary perspectives, and transferable indicators across assessment programs.

From an operational perspective, the development of context questionnaire items for use with students in LMICs raises specific challenges related to comprehensibility and appropriateness of wording. Questionnaire items need to be formulated in a manner that is as simple and concrete as possible so that they can be understood by young respondents across diverse linguistic and cultural settings. Different international and regional assessments, accordingly, follow distinct traditions and styles of operationalization.

Finally, while psychometric information from previous assessment cycles can provide useful insights into response distributions and measurement limitations, such evidence is often incomplete, context-dependent, and not directly transferable to revised or newly adapted items. Overreliance on statistical associations with achievement risks excluding indicators that are of high policy relevance but show limited variation or weak correlations in specific datasets. There is therefore a need for an approach that balances conceptual relevance, policy usefulness, qualitative judgment, and available quantitative evidence without overstating the evidentiary status of any single source.

3.2 Aims of the Study and Research Questions

ILSAs, including those conducted by the IEA, play a central role in generating comparable evidence on learning outcomes and their associated context factors. However, ensuring that context questionnaires adequately capture the conditions of learning in low- and lower-middle-income countries (LMICs) at the end of primary education remains a key challenge.

This study accordingly aimed to identify and operationalize context indicators that are relevant for understanding learning outcomes and education systems at the end of primary education in LMICs. The primary objective was to support the development and refinement of context questionnaires in international and regional learning assessments by providing a structured, evidence-informed set of indicators and illustrative questionnaire operationalizations. While end-of-primary IEA assessment frameworks served as the main operational reference, the study

was explicitly designed to inform broader applications, including regional assessments, national monitoring systems, and international reporting initiatives.

The research questions were organized to reflect a top-down analytical approach. They began with a comparison of international and regional assessment frameworks to identify common domains relevant in LMIC contexts, proceeded with the identification and justification of indicators (or constructs), and then addressed the operationalization of these indicators using existing questionnaire items. The final research question focused on qualitative country-level feedback to assess the clarity (degree to which the question was assumed to be understandable as intended), differentiation potential (degree to which the item was assumed to distinguish between respondents), and policy relevance (degree to which the information was relevant for policy or system monitoring) of proposed questionnaire operationalizations. Together, these questions defined the analytical scope of the study and guided the development of user-oriented outputs.

Overall research question

How can robust, policy-relevant context indicators for learning at the end of primary education be identified and operationalized to better reflect the educational realities of LMICs and to inform the development of context questionnaires in LaNA and other learning assessments?

RQ1: Framework comparison

- What commonalities and differences exist between IEA end-of-primary context frameworks and regional assessment frameworks (ERCE, PASEC, SACMEQ, SEA-PLM, PISA-D, AMPL), and how can topics and operationalizations from regional and LMIC-focused assessments strengthen IEA-based context frameworks to better reflect LMIC educational realities while maintaining cross-national comparability?

RQ2: Indicator identification and justification

- Which background factors emerge as robust and policy-relevant context indicators for LMICs, based on their documented associations with achievement outcomes in core subjects such as mathematics and reading, or their relevance for education policy and system monitoring as evidenced in existing literature and prior assessment experience?

RQ3: Operationalization

- How can the identified indicators be operationalized using existing questionnaire frameworks and items, anchored in IEA studies and complemented by regional and LMIC-focused initiatives where appropriate, to inform the development of context questionnaires in LaNA and other learning assessments?

RQ4: Integrated Item Evaluation and Item Suitability Assessment

- How can the suitability of questionnaire items for LMIC contexts be systematically evaluated and synthesized into a guidance framework?
- How do LaNA National Research Coordinators assess item clarity, differentiation, and policy relevance?
- How do items perform empirically in terms of psychometric characteristics?

- How can expert and empirical evidence be integrated into an item suitability classification and final guidance rating?

4 Literature and Framework Review

This chapter synthesizes findings from the systematic literature review and the review of existing international and regional assessment frameworks to identify educational and contextual factors that are particularly relevant for understanding learning outcomes in LMICs. The review focused on factors that have been shown to be associated with achievement in reading and mathematics at the end of primary education and that are, in principle, measurable through large-scale context questionnaires.

Rather than aiming to provide an exhaustive account of all determinants of learning, the review adopted a targeted perspective aligned with the objectives of this project: to inform the enhancement of IEA context questionnaires for use in LMICs. Accordingly, the chapter emphasizes factors that (a) recur across empirical studies and regional assessments, (b) show stronger or more variable associations with achievement in LMICs than in high-income countries, and (c) can plausibly be captured through student, teacher, or school questionnaires administered at scale.

Section 4.1 provides an overview of educational and contextual factors identified in the literature as particularly relevant in LMICs. The review is organized across three analytical levels. Section 4.1.1 focuses on student-level factors, including home learning environments and their role in language and literacy development, as well as patterns of absenteeism, grade repetition, and dropout in relation to health conditions, economic inequality, and parental involvement.

Section 4.1.2 then examines teacher- and classroom-level factors, synthesizing evidence on instructional practices, classroom organization, teacher preparation, and working conditions that influence learning outcomes in primary education.

Section 4.1.3 focuses on school-level factors, including basic infrastructure and services (such as electricity, water, sanitation, and internet access), the availability and quality of textbooks and other learning materials, core school effectiveness characteristics (including curriculum implementation, teaching time, and school climate), and teacher qualifications and professional development opportunities. Emphasis is placed on the variability and conceptual complexity involved in defining and measuring teacher qualifications across LMIC contexts.

Finally, Section 4.1.4 discusses education-related SDG indicators, identifying research that has examined or used ILSA context questionnaire data, highlighting relevant indicators beyond traditional socioeconomic measures, and assessing gaps in current IEA instruments with respect to SDG monitoring in LMICs.

In addition to the academic literature identified through the systematic review, the chapter also draws on literature referenced in the questionnaire frameworks of selected regional and international large-scale assessments. These sources were reviewed to ensure that the factors discussed reflect the empirical research underpinning existing regional assessment instruments in LMIC contexts. This complementary review supports alignment between the literature-based evidence and the constructs operationalized in regional assessment frameworks and helps to identify areas of convergence as well as gaps in current questionnaire coverage.

4.1 Educational and Contextual Factors in LMICs

There is a substantial body of research examining contextual factors that influence educational outcomes. However, findings derived primarily from studies conducted in high- and upper-middle-income country (HIC /UMIC) contexts do not always translate directly to those in LMICs. Differences in resource availability, institutional capacity, and baseline learning conditions mean that both the relevance and the explanatory power of commonly studied factors can vary considerably across contexts.

As noted by Hattie (Hattie, 2009) in his synthesis of meta-analyses on student achievement, “the situation is quite different in less resourced nations (e.g., throughout Africa).” In this discussion, Hattie refers to differences in the variability and relative importance of school-level inputs. In less-resourced systems, schools tend to account for a larger share of the variance in student achievement than is typically observed in high-income settings. Where basic conditions for teaching and learning are not universally met, variation in school resources, organization, and instructional conditions becomes a more salient determinant of learning outcomes. (Mejía-Rodríguez & Kyriakides, 2022), for example, show that the effects of contextual and system-level factors on student achievement are not uniform, but vary in both magnitude and relevance depending on broader societal and institutional conditions, such as levels of economic development, resource availability, and education system capacity.

Other research similarly indicates that factors often treated as background conditions in studies focusing on HIC and UMIC systems remain unevenly distributed across many LMIC education systems and may therefore exert a stronger influence on student learning. This includes factors like access to basic infrastructure, instructional materials, qualified teachers, and sufficient teaching time (Friedman, Schwantner, Spink, Tabata, & Waters, 2016; Hattie, 2009; Heyneman & Lee, 2013).

LMICs should not, however, be treated as a homogeneous group. Literature consistently reflects substantial heterogeneity both within and across LMICs with respect to school quality, governance arrangements, and socioeconomic conditions. Comparative research shows that the magnitude and relevance of educational inputs vary across systems depending on broader contextual factors such as levels of economic development, resource availability, and institutional capacity (Mejía-Rodríguez & Kyriakides, 2022; World Bank, 2018).

Regionally, evidence from Eastern and Southern Africa shows wide variation in access to basic infrastructure, teacher availability, and learning conditions across countries and between urban and rural areas (Friedman et al., 2016). Similarly, in Latin America, ERCE countries exhibit substantial heterogeneity in school resources, instructional practices, and socioeconomic inequalities across education systems (Sandoval-Hernández et al., 2019). Important differences also exist between regional assessment programs. Compared to PASEC and SACMEQ, ERCE includes a larger share of high-income and upper-middle-income countries (see Table 2-1). These differences in country composition and measurement needs reflect broader regional disparities in resource availability, system capacity, and patterns of inequality, which in turn shape how educational inputs relate to student learning outcomes (Mejía-Rodríguez & Kyriakides, 2022; Okubo, 2024).

4.1.1 Student-Level Factors

ILSAs frequently include indicators capturing students’ home learning environments, patterns of school participation, and broader socioeconomic constraints (Brese & Mirazchiyski, 2013).

These factors provide important context for interpreting differences in student achievement across education systems.

In many LMIC education systems, disparities in household resources, health conditions, and access to educational support outside the classroom remain substantial. Such differences can influence students' opportunities to engage with learning activities both within and beyond the school environment. Student-level indicators are therefore particularly important for understanding variation in learning outcomes and for interpreting assessment results in diverse educational contexts (Friedman et al., 2016; Mejía-Rodríguez & Kyriakides, 2022; Sandoval-Hernández et al., 2019). The following subsections examine several student-level factors that have been identified in the literature as particularly relevant for understanding learning outcomes in LMIC contexts.

The indicators reviewed in this study are intended to capture factors that may be relevant across a broad range of LMIC education systems. Nevertheless, their relevance and measurement properties cannot be assumed to be identical in every national or regional context. Additional context-specific indicators may therefore be needed to capture locally important conditions, practices, or constraints. This consideration is particularly important for home-possession measures, because the prevalence and socioeconomic meaning of individual possessions can vary across countries and over time, potentially limiting the comparability and local validity of a common item set (Rutkowski & Rutkowski, 2013).

4.1.1.1 Home Learning Environments and Language Exposure

A consistently identified predictor of student achievement in ILSAs is the home learning environment, typically operationalized through indicators of socioeconomic status (SES) such as parental education, the number of books in the home, the availability of educational resources, and opportunities for literacy-related activities (Brese & Mirazchiyski, 2013). These indicators have shown robust associations with student performance across diverse education systems (Hattie, 2009; Sandoval-Hernández et al., 2019).

As discussed in Section 4.1, LMIC education systems are characterized by substantial heterogeneity, and differences in household resources and living conditions are an important mechanism through which this heterogeneity translates into unequal learning opportunities. Two features generally distinguish LMIC contexts from HIC settings. First, underlying socioeconomic inequality within countries tends to be more pronounced (UNESCO, 2020; World Bank, 2018), producing wider dispersion in the material and cultural resources available to students. Second, absolute levels of household resources are markedly lower: substantial proportions of students in LMICs grow up without access to books, quiet study spaces, electricity, or other basic learning supports (UNESCO, 2020). This combination of greater dispersion and lower baseline resources also has methodological implications, as ILSA socioeconomic background scales originally developed for HIC contexts can exhibit floor effects in LMICs and may not fully discriminate among the most disadvantaged students (Rutkowski & Rutkowski, 2013).

These patterns are mirrored at the school level, where shortages of instructional materials, infrastructure, and qualified teachers remain more prevalent than in high-income systems (Fuller, 1987; Scheerens, 2000; World Bank, 2018). The compounding of resource deprivation at home and at school is central to understanding the social gradient in achievement in LMICs. The Heyneman–Loxley effect (Heyneman & Loxley, 1983) provides a useful comparative framework to understand this reality: in lower-income settings, variation in school and material

resources was originally found to play a larger role in shaping learning outcomes than family background, whereas in more affluent systems family background tended to dominate. Subsequent research has questioned (Baker, David, P., Goesling, & Letendre, Gerald, K., 2002) and partially reaffirmed (Chudgar & Luschei, 2009) the original claim, suggesting that the effect is conditional on national income and the level of within-country inequality. While the empirical strength of the effect remains debated, its core intuition continues to inform the present analysis: in LMICs, both the absolute scarcity of basic learning resources and the wider dispersion of household and school conditions are likely to translate more directly into achievement gaps based on the home background than in higher-income contexts, making socioeconomic background indicators particularly important for interpreting differences in learning outcomes.

Homework practice represents another important factor of learning beyond the classroom. The frequency and nature of homework assignments, as well as the support available to complete them, can influence students' opportunities to practice and consolidate newly acquired skills, although their effects vary across contexts and may depend on home environments and levels of support (Cooper, Robinson, & Patall, 2006; Hattie, 2009). In LMIC contexts, where instructional time is often constrained, homework may play a more prominent role in supporting learning by compensating for limited time on task in classrooms (Abadzi, 2007). Emerging evidence from cross-country studies further suggests that homework may be more consistently associated with positive learning outcomes in LMIC countries than in high-income contexts, although these effects remain contingent on students' access to support and resources at home (Flannery, Palcic, & Clavel, 2024).

Parental involvement in children's learning, including support with schoolwork and engagement with school-related activities, has also been positively associated with student achievement. However, empirical findings are mixed, as higher levels of parental support may both promote learning and reflect compensatory responses to students' academic difficulties. The extent and nature of such involvement varies across socioeconomic contexts, particularly when parents have limited time, educational background, or resources to support learning at home (Desforges & Abouchaar, 2003; Hattie, 2009). Meta-analytic evidence suggests that the relationship between parental involvement and achievement is positive but heterogeneous, with effects varying depending on socioeconomic context and the type of involvement (Chudgar & Luschei, 2009; Joshua Jeong; Tan, Lyu, & Peng, 2020).

Indicators such as housing quality (e.g., construction materials), access to basic services (including electricity, safe drinking water, and sanitation), and household composition (e.g., number of people living in the household) are commonly used as proxy measures of socioeconomic background in ILSAs. These indicators are particularly important in contexts where direct measures of income are difficult to obtain and where material living conditions vary substantially across households (Brese & Mirazchiyski, 2013; UNESCO, 2020). Such factors influence students' opportunities to study, their health and well-being, and the overall conditions under which home learning takes place.

Language exposure within the home environment represents another important factor shaping learning in LMICs. In many LMIC education systems, students enter school speaking a home language that differs from the official language used for instruction. Research suggests that such language mismatches can affect reading comprehension, participation in classroom activities, and overall academic performance, particularly during the early years of schooling (Hochweber & Vieluf, 2018). Similar challenges regarding language used in the home versus the

testing language were reported by National Research Coordinators (NRCs) across participating LaNA 2023 Linking Study countries. Indicators capturing home language use and language of instruction therefore provide important contextual information for interpreting achievement data in linguistically diverse education systems. Understanding the linguistic environments in which students develop literacy skills can help explain variation in reading outcomes across students and schools, particularly in contexts where linguistic diversity exists alongside broader socioeconomic inequalities.

4.1.1.2 Student Participation, Engagement, and Educational Trajectories

Patterns of school participation represent another important dimension of student-level variation. Regular attendance at school is a fundamental condition for sustained learning and favorable educational outcomes; research has shown that student absenteeism is associated with a range of academic and behavioral risk factors, including lower levels of academic engagement and an increased likelihood of dropout (Gubbels, van der Put, & Assink, 2019). While the causes of absenteeism are multifaceted, consistent participation in schooling remains an essential prerequisite for learning progress.

In many LMIC contexts, patterns of school participation may be influenced by both school-related conditions and broader household circumstances. Economic pressures, household responsibilities, transportation barriers, and the quality of school environments can all contribute to irregular attendance or interruptions in schooling (Friedman et al., 2016).

Students' prior educational experiences and trajectories can also help explain differences in later learning outcomes. Participation in early childhood education, including its duration and quality, can support school readiness and the development of foundational cognitive and socioemotional skills, although the benefits depend substantially on the quality and context of the program in question (Bendini & Devercelli, 2022). Delayed school entry, interruptions in schooling, and grade repetition may indicate reduced opportunities to learn or accumulated learning difficulties. Grade repetition is therefore an important indicator of students' educational trajectories, although evidence regarding its causal effects on subsequent achievement is mixed and varies across education systems and student groups (Goos, Pipa, & Peixoto, 2021).

Longer or more burdensome commutes to school are likewise usually associated with lower learning outcomes, as they are associated with lost time, fatigue, absenteeism, and poorer mental health. Moderate-length, physically active commutes may be beneficial, but very long commutes—especially on foot or bus—are generally found to undermine learning (David, Temba, & Matondo, 2025).

4.1.1.3 Health, Nutrition, and Socioeconomic Constraints

A growing body of research supports the role of health, nutrition, and broader socioeconomic conditions in shaping the opportunities students have to learn. Health problems and nutritional deficiencies can affect cognitive development, concentration, and school participation, particularly among younger students. In addition to physical health, students' mental and socio-emotional well-being, including anxiety, emotional distress, sense of belonging, motivation, self-confidence, and life satisfaction, is associated with differences in concentration, participation, and engagement in learning activities.

These factors are increasingly recognized as important influences on learning, particularly in contexts where students may face multiple socioeconomic stressors (OECD; United Nations

Children’s Fund, 2021). In contexts where access to healthcare and adequate nutrition remains uneven, these factors may have a substantial influence on students’ ability to benefit from schooling (Friedman et al., 2016).

Socioeconomic constraints can also affect students’ engagement with education in other ways. Household poverty and domestic responsibilities, as well as the need for children to engage in income-generating activities or labor outside the household, may reduce the time available for study and increase the likelihood of irregular attendance or leaving school early. Research from the ILO and UNICEF indicates that children involved in child labor are less likely to attend school regularly and are at greater risk of falling behind or dropping out altogether, particularly in low-income contexts (International Labour Office & United Nations Children’s Fund, 2021). Such conditions can create additional barriers to sustained attendance and participation in school, contributing to disparities in learning outcomes across socioeconomic groups.

Context questionnaires in ILSAs therefore frequently include indicators capturing students’ socioeconomic background and learning environments outside school (Brese & Mirazchiyski, 2013). These indicators help provide a more comprehensive understanding of the conditions under which students learn, allowing researchers to interpret achievement differences in relation to broader social and economic contexts (Mejía-Rodríguez & Kyriakides, 2022; Sandoval-Hernández et al., 2019).

4.1.2 Teacher and Classroom-Level Factors

Teacher- and classroom-level factors determine how teaching and learning take place day-to-day. These factors represent the immediate interface between institutional resources and students’ learning experiences. Differences in teacher preparation, instructional practices, employment conditions, and classroom organization can therefore play important roles in shaping educational outcomes.

Variation in teacher-level conditions may be particularly pronounced in LMIC education systems. Rapid expansion of school enrollment in many countries has required large-scale recruitment of teachers through diverse certification pathways and employment arrangements (Evans & Yuan, 2018; Singh, 2021). As a result, teachers within and across schools may differ considerably in terms of training, teaching responsibilities, and working conditions.

ILSAs frequently capture such variation through context questionnaire indicators related to teacher preparation, instructional confidence, classroom practices, and professional working conditions. These indicators provide insights into how teaching processes operate within schools and how they may contribute to differences in student learning outcomes.

4.1.2.1 *Teacher Preparation and Professional Background*

At the teacher level, professional background refers not only to formal qualifications, discussed in the subsequent section at the school level, but also to how teachers are assigned, positioned, and supported in their day-to-day work. Indicators such as the grades or classes taught, the duration and nature of pre-service preparation, certification pathways, and employment arrangements provide information on the match between teachers’ preparation and their classroom responsibilities.

In many LMIC contexts, these dynamics result in substantial variation in teachers’ professional backgrounds within schools. Teachers may differ in training, pedagogical preparation, and

employment stability, which can affect continuity of instruction, subject specialization, and the distribution of teaching responsibilities (Hanushek, 1995; Scheerens, 2004).

The grades or classes assigned to teachers also impact instruction. Teaching in the early grades (particularly reading and foundational numeracy), for example, may require specialized pedagogical approaches. Mismatches between teachers' preparation and their assigned grade levels may therefore affect instructional effectiveness. Indicators capturing grade-level teaching assignments can thus provide valuable information for interpreting achievement patterns. In addition to grade-level assignments, subject-matter assignments are important when considering alignment between teachers' preparation and the subjects they teach. Such alignment may be particularly impactful in contexts where teachers are required to teach multiple subjects or are assigned outside their area of specialization, as is often the case in resource-constrained education systems (Friedman et al., 2016; Mejía-Rodríguez & Kyriakides, 2022; Scheerens, 2004).

Employment conditions such as contract type and professional position within the school may also influence teacher stability and working conditions. In systems where temporary or contract-based appointments are common, employment insecurity may affect professional commitment and workload, with potential implications for classroom practices (Okubo, 2024; Scheerens, 2000).

Finally, the prevalence of additional paid employment outside the school system is frequently discussed in LMIC education research, particularly in contexts where teacher salaries remain relatively low. Holding additional jobs may reduce the time available for lesson preparation, professional development, or engagement with students, thereby influencing instructional quality (Friedman et al., 2016).

4.1.2.2 Instructional Confidence, Classroom Practice, and Language of Instruction

Teacher-level factors related to instructional confidence and classroom practice play an important role in shaping learning outcomes. The perceived preparedness and confidence of teachers in their teaching can influence factors like classroom organization, instructional clarity, and responsiveness to student needs. While such constructs are frequently examined in HIC/UMIC education systems, they may be particularly relevant in LMIC contexts, where teachers often work under resource constraints and manage classrooms with diverse learning needs (Li, 2021; Scheerens, 2004).

The way that teaching and learning resources are *used* in the classroom also influences how strongly their availability translates into improved student learning outcomes. As discussed in below in section 4.1.3, the presence of textbooks or instructional materials does not automatically lead to improved learning outcomes; their impact depends largely on how they are integrated into classroom instruction (Hanushek, 1995; Wagemaker, 2020b). In LMIC contexts, differences in teachers' capacity to adapt instructional materials, manage large or multi-grade classrooms, and compensate for limited resources may substantially shape students' learning experiences.

Language of instruction represents another important classroom-level factor. In many LMIC education systems, teachers work with students whose home language differs from the language used for instruction. This often reflects historical and policy-related factors, such as the adoption of former colonial or national official languages as media of instruction in linguistically diverse societies (UNESCO, 2025). Such language differences can influence

reading comprehension, participation in classroom activities, and overall academic performance, particularly in the early years of schooling (Friedman et al., 2016; Hochweber & Vieluf, 2018).

4.1.2.3 Teacher Attendance and Work Conditions

Teacher attendance and working conditions represent another important dimension of classroom-level variation in LMIC education systems. In settings where teacher shortages, large class sizes, and resource constraints are common, the consistent presence of teachers in the classroom is a fundamental condition for maintaining instructional continuity. Teacher absenteeism has been widely documented in LMIC contexts; for example, Chaudhury et al. found that a substantial share of teachers may be absent on a given school day, reducing time for instruction (Chaudhury, Hammer, Kremer, Muralidharan, & Rogers, 2006). This contributes to reduced opportunities for learning and, in turn, disruptions in curriculum coverage (Friedman et al., 2016; Liu, Zhang, & Hu, 2024; Zhang & Holden, 2023).

Beyond its immediate impact on teaching time, teacher absenteeism is also linked to more wide-spread patterns of instructional quality and teacher engagement. Evidence from LMIC contexts shows that high rates of teacher absence substantially reduce effective instructional time and limit students' exposure to teaching, with negative consequences for learning outcomes (Abadzi, 2007; Bold et al., 2017; Karamperidou, Brossard, Peirolo, & Richardson, 2020). Teacher absenteeism is also associated with factors such as low motivation, weak accountability, and poor working conditions, particularly in disadvantaged or remote schools (Guerrero et al., 2012). These patterns suggest that absenteeism reflects not only time loss, but also deeper challenges related to teacher engagement and the functioning of education systems.

Teacher attendance patterns may also be influenced by employment conditions and workload structures. In some LMIC contexts, teachers engage in additional paid employment outside the school system. Such arrangements may reduce the time available for lesson preparation, professional development, or engagement with students, potentially affecting instructional quality (Friedman et al., 2016). Teacher job satisfaction represents an important aspect of teachers' working conditions. Evidence from LMIC contexts shows that it is strongly shaped by factors such as remuneration, workload, school leadership, and resource availability (Sahito & Väisänen, 2019); broader (not LMIC-specific) research indicates that job satisfaction is associated with teacher motivation, retention, and instructional processes (Wartenberg, Aldrup, Grund, & Klusmann, 2023).

Finally, differences in contract stability, supervision arrangements, and accountability mechanisms may further influence teacher attendance and professional commitment across schools (Okubo, 2024; Scheerens, 2000).

4.1.2.4 Supervision and Instructional Oversight

External supervision and instructional oversight are other important aspects of the environment in which teachers work. In many education systems, mechanisms such as school inspections, supervisory visits, or other forms of external monitoring are intended to support instructional quality, promote accountability, and ensure that teaching aligns with national curriculum standards.

The presence and effectiveness of these oversight mechanisms can vary considerably within and among LMICs. Differences in administrative capacity, resource availability, and governance

structures may influence both how frequently supervision takes place and how meaningful it is for teachers. Teachers may therefore receive very different levels of professional guidance, feedback, and monitoring depending on the school or country in which they teach.

Research on school effectiveness has long highlighted the role of instructional leadership and accountability structures in supporting effective teaching practices (Scheerens, 2000; Scheerens, 2004). When supervision is regular and constructive, it can help teachers refine instructional strategies, align classroom practices with curricular expectations, and respond to challenges that arise in the classroom.

4.1.3 School-Level Factors

A substantial body of research concerns the role of school resources, organizational structures, and institutional practices in shaping the environments in which teaching and learning take place. While many of these factors are relevant across education systems, their influence may be particularly pronounced in LMIC contexts (Hattie, 2009; Mejía-Rodríguez & Kyriakides, 2022; Scheerens, 2004), where variation in basic resources and school organization is considerable.

Schools differ markedly in terms of infrastructure, availability of learning materials, staffing arrangements, and organizational practices in many LMIC education systems. These differences can directly affect the quality of instruction, as well as students' opportunities to learn. Staffing levels, including the number of teachers and support staff available within a school, also influence class size, workload distribution, and the school's capacity to deliver instruction effectively. As a result, indicators capturing school-level conditions are widely included in context questionnaires administered as part of ILSAs (Brese & Mirazchiyski, 2013).

The following sections review key school-level factors that have been identified in the literature as particularly relevant in LMIC contexts. These include basic infrastructure and services, the availability and quality of instructional materials, organizational characteristics related to school effectiveness, and the qualifications and professional development opportunities available to teachers.

4.1.3.1 *Infrastructure (Electricity, Water, Internet, Sanitation)*

A central theme in research on education in LMICs is the role of basic school infrastructure in enabling effective teaching and learning. Access to electricity, safe drinking water, adequate sanitation facilities, and basic connectivity such as internet access is uneven across and within many LMIC education systems (Friedman et al., 2016; Mejía-Rodríguez & Kyriakides, 2022). Where such infrastructure is lacking, instructional time may be reduced overall, learning environments may become less conducive to sustained engagement, and student health and attendance may be negatively affected (Friedman et al., 2016; Wang & Fawzi, 2020).

The physical condition of classroom facilities, including adequate space, desks, lighting, and ventilation, also shapes students' opportunities to learn. Overcrowded classrooms and insufficient facilities remain common in LMICs, particularly in rural or resource-constrained areas, potentially limiting effective teaching and learning (Friedman et al., 2016; Scheerens, 2004).

School-based nutrition programs and the provision of meals represent another important factor influencing learning conditions in some LMIC settings. Research suggests that initiatives to offer meals at school can improve attendance, concentration, and learning outcomes (although the

evidence is less consistent). This is particularly true for socioeconomically disadvantaged students (Wang & Fawzi, 2020).

4.1.3.2 Availability and Quality of Learning Materials

Closely related to school infrastructure is the availability and quality of instructional materials, particularly textbooks and other subject-specific learning resources. A substantial body of research documents persistent shortages of textbooks in many LMIC education systems, as well as high student-to-book ratios, especially in foundational subjects such as reading and mathematics (Friedman et al., 2016; Zhang & Holden, 2023). Unlike in many HIC/UMIC contexts, where textbook provision is nearly universal, access to learning materials continues to represent an important source of inequality across schools in LMIC settings (Hanushek, 1995). More recently, ERCE 2019 highlighted substantial disparities in the availability of instructional resources across schools and education systems in Latin America and the Caribbean (UNESCO, 2020). Access to digital devices and technology resources has also become an increasingly important dimension of the learning environment; as with other school-level factors, there are significant disparities in both availability and effective use across many LMIC contexts (UNESCO, 2023).

Regional assessments frequently operationalize this construct through relatively straightforward and observable indicators, such as whether students have access to their own textbooks or whether schools maintain functional libraries or reading corners. Evidence from both empirical studies and assessment practice suggests that such indicators are feasible to collect and can be meaningfully associated with learning outcomes in LMIC contexts (Friedman et al., 2016; Sandoval-Hernández et al., 2019). Research on the design of ILSA context questionnaires echoes the importance of including clearly defined and observable school resource indicators in order to support reliable cross-national comparisons (Wagemaker, 2020b).

The literature consistently cautions against assuming that the presence of specific resources, practices, or conditions will automatically lead to improved learning outcomes. Their effects depend on how they are implemented, experienced, and combined with other features of the learning environment. In the case of instructional materials, for example, effectiveness depends not only on availability but also on how materials are integrated into classroom instruction and pedagogical practice (Hanushek, 1995; Scheerens, 2004). Indicators capturing both the availability and the use of learning materials are therefore important for understanding variation in student achievement.

4.1.3.3 School Effectiveness Factors (Curriculum Quality, Teaching Time, School Climate, Orderly Environment)

In addition to material inputs, the literature identifies several organizational and process-related school factors that influence learning outcomes in LMIC contexts. Curriculum implementation, the effective use of instructional time, school climate, and the presence of an orderly and safe learning environment (Hattie, 2009; Scheerens, 2000; Scheerens, 2004), for example, are all referenced. While these constructs are also commonly examined in studies focusing on HIC and UMIC education systems, their operationalization and relative importance often differ in LMIC settings.

Disruptions to instructional time due to teacher absenteeism, school closures, overcrowding, or weak school management are frequently documented in LMIC contexts and may constrain students' opportunities to learn (Friedman et al., 2016; Zhang & Holden, 2023). Indicators

capturing basic aspects of school climate—such as safety, discipline, and orderliness—have also been identified as important factors that shape learning environments in many LMIC education systems (Heyneman & Lee, 2013).

How schools are structured and governed further shapes instructional conditions and resource allocation. School ownership and governance arrangements—for example, whether a school is public, private, or faith-based—are often associated with differences in funding, accountability mechanisms, and levels of autonomy. These institutional differences can, in turn, influence the availability of resources, staffing patterns, and instructional practices across schools, particularly in LMIC contexts. Examining such characteristics therefore provides important context for interpreting variation in student achievement (Heyneman & Lee, 2013; Mejía-Rodríguez & Kyriakides, 2022; Okubo, 2024). In addition, characteristics of school leadership such as the age and gender of the school principal are included in some questionnaires as contextual variables. While these characteristics are not consistently linked to student achievement, they may provide useful information for understanding variation in leadership experience, management practices, and school organization across education systems.

The configuration and type of school, e.g., the grades covered, the presence of multi-grade arrangements, or the existence of different language streams, also shape teaching practices and classroom organization. Multi-grade teaching contexts and combined primary and secondary institutions remain common in some LMIC systems, particularly in rural or resource-constrained areas. These arrangements can affect instructional time, curriculum coverage, and teacher workload. Questions related to school type therefore provide important contextual information for interpreting student achievement data in LMIC contexts (Friedman et al., 2016; Scheerens, 2004).

Organizational arrangements such as double- or multi-shift schooling are used in some LMIC systems to accommodate high student enrollment. While these approaches help expand access to education, they can also reduce effective instructional time and limit opportunities for extracurricular or remedial support, impacting learning conditions (Friedman et al., 2016; Zhang & Holden, 2023).

School-level policies regarding grade retention and student advancement also shape students' learning trajectories, particularly in systems where repetition rates remain relatively high (Friedman et al., 2016; Okubo, 2024).

Finally, the availability of structured academic support, such as remedial instruction or tutoring in reading and mathematics, can influence how schools respond to foundational learning gaps, especially where early skill deficits are widespread (Mejía-Rodríguez & Kyriakides, 2022; Scheerens, 2004).

4.1.3.4 Teaching Workforce and Professional Development

Teacher qualifications and opportunities for professional development constitute another important dimension of school capacity in LMIC contexts. While teacher-related indicators are widely included in international context questionnaires, the meaning and comparability of formal qualifications can vary substantially across education systems. Differences in training duration, certification pathways, and the quality of teacher education institutions mean that similar qualification labels may reflect different levels of pedagogical preparation (Li, 2021; Mejía-Rodríguez & Kyriakides, 2022).

Empirical findings regarding the relationship between formal teacher qualifications and student achievement are mixed. In contexts where teachers' qualifications vary widely, as is the case in many LMICs (World Bank Group, 2026), minimum training thresholds may have a meaningful impact on instructional quality. However, additional years of formal education do not consistently translate into improved learning outcomes (Hanushek, 1995; Scheerens, 2004). It becomes therefore important to distinguish between formal certification and actual instructional competence. Evidence from several LMIC contexts, reported by the World Bank, further suggests that formal qualifications do not always correspond to teachers' subject knowledge or instructional effectiveness. Qualification indicators should therefore be complemented by measures of actual teaching practice and professional support (World Bank Group, 2026).

Recent research and assessment frameworks increasingly emphasize indicators related to *in-service* professional development and *ongoing* training opportunities. Access to professional development can help teachers strengthen subject knowledge, improve pedagogical practices, and adapt instruction to diverse classroom contexts. However, opportunities for training and professional support may vary substantially across schools, and vary based on factors like leadership practices, resource availability, and national policy priorities (Friedman et al., 2016; Wagemaker, 2020b).

Context questionnaires should therefore capture not only teachers' formal qualification levels but also their access to professional development (and other forms of instructional support). Such indicators are particularly relevant in LMIC education systems, where rapid teacher workforce expansion, decentralized certification pathways, and uneven training provision may lead to substantial variations in teacher preparation across schools.

4.1.4 SDG-Related Information Needs

Sustainable Development Goal (SDG) indicators serve as internationally agreed, measurable benchmarks for tracking progress on critical global outcomes, including education quality, poverty reduction, and infrastructure (D'Adamo, Della Sciucca, Gastaldi, & Lupi, 2025; Sachs, Lafortune, Fuller, & Iablonski, 2025). Empirical evidence shows that LMICs lag behind in performance across multiple goals compared to their high-income peers, which is the result of enduring structural constraints such as limited resources, weaker institutional capacities, and political instability that disproportionately affect these regions (Confraria, Ciarli, & Noyons, 2024; D'Adamo et al., 2025).

Data gaps and statistical capacity limitations are notably more prevalent in low-income contexts, which creates significant obstacles to building an inclusive evaluation of the 2030 Agenda and undermines evidence-based policy planning (D'Adamo et al., 2025). Integrating SDG indicators into context or background questionnaires within ILSAs not only aligns data collection with global priorities but also provides a targeted assessment of the contextual constraints and progression toward universal goals within these specific settings (UIS, 2025).

ILSAs offer several important advantages for monitoring selected indicators of the (SDGs), especially through their standardized assessment design, cross-national comparability, and the systematic collection of contextual information via background questionnaires. As stated by the UNESCO Institute for Statistics (UIS), ILSAs are not intended to provide comprehensive coverage of the SDG agenda (UIS, 2025). Nevertheless, a limited subset of SDG indicators, particularly those related to learning conditions, equity, educational resources, and participation, can be meaningfully approximated using data derived from ILSA context

questionnaires. At the same time the report states that ILSA context questionnaire reporting faces important limitations, including the indirect nature of many measures, the exclusion of out-of-school and marginalized populations, the reliance on self-reported information, and the fact that a substantial proportion of SDG indicators require administrative, longitudinal, or system-level data that lie beyond the scope of ILSAs.

Nevertheless, UIS (2025) identified several SDG-related indicators as either measurable or partially measurable using ILSA context data. Building on this classification, the present project systematically mapped those SDG indicators identified by UIS as measurable against indicators derived from the comparative review of the LMICs targeting questionnaire instruments. This alignment subsequently informed the project's rating scheme, within which explicit SDG relevance constituted one criterion for retaining indicators that are regarded as specifically relevant in low- and lower-middle-income contexts.

Specifically, indicators involving school infrastructure, teacher qualifications, home environments, and ICT usage can be at least partially tracked, though the UIS paper states that they may lack the system-level coverage or administrative reporting detail (such as student-to-teacher ratios or specific legal guarantees) required for full SDG compliance.

Considering the importance of the SDG agenda and the recognized data gaps in LMICs, and despite the limitations outlined above, this project builds on the potential of ILSAs to generate proxy and supplementary evidence for monitoring selected SDG indicators.

Table 4-1 provides an overview of the SDG indicators included in this research project.

Table 4-1: SDG indicators measurable via ILSA context questionnaires according to UIS (2025)

Indicator ID	Description	ILSA Sources	Measurability & Limitations
4.1.5	Children over-age for grade	PISA, TIMSS, PIRLS, ERCE	Partial. Identifies entry age and repetition but doesn't track longitudinal progression.
4.2.1	Developmentally on track (under 5s)	PISA, ERCE, TIMSS, PIRLS	Limited. Measures access to early learning but not direct developmental outcomes.
4.2.2	Organized learning participation	PISA, ERCE, PIRLS, TIMSS, PILNA	Partial. Tracks age of entry and attendance but lacks data on program quality or duration.
4.2.3	Stimulating home environments	ERCE, PISA, PIRLS, TIMSS, SEA-PLM	Partial. Measures frequency of activities (reading, singing) but relies on subjective self-reporting.
4.2.4	ECE enrollment rate	PISA, PIRLS, TIMSS	Partial. Captures attendance but lacks disaggregated population data needed for full rates.
4.4.1	ICT skills by type	ICILS, PISA, ERCE, TIMSS	Partial. Measures frequency of tool usage (spreadsheets, etc.) rather than actual proficiency or mastery.

Indicator ID	Description	ILSA Sources	Measurability & Limitations
4.4.2	Digital literacy proficiency	PISA, ICILS	Partial. Self-reported data lacks direct assessment of problem-solving in tech environments.
4.4.3	Educational attainment rates	PIRLS, PISA	Limited. Focuses on parent/family structure; needs integration with national census data for full population rates.
4.5.2	Home language instruction	PILNA, PISA, TIMSS, ERCE	Partial. Captures home language but requires external sources to confirm the actual language of instruction.
4.a.1	Basic school services/infrastructure	ERCE, PISA, PASEC, SEA-PLM	Partial. Effective for electricity/water/tech; often misses data on adapted materials for disabilities.
4.a.2	Experience of bullying	PISA, TIMSS, SEA-PLM, PILNA	Partial. High reliance on self-reporting introduces subjectivity and social desirability bias.
4.c.1	Teacher qualifications	PASEC, TIMSS, ERCE, PISA, SEA-PLM, PILNA	Partial. Tracks certificates and training duration but may not align with specific national standards.
4.c.2	Pupil-trained teacher ratio	ERCE, TIMSS, PISA	Partial. ILSAs measure qualifications but often lack the simultaneous student enrollment counts to form a ratio.
4.c.4	Pupil-Qualified Teacher Ratio	ERCE, TIMSS	Partial. ILSAs measure qualifications but often lack the simultaneous student enrollment counts to form a ratio.
4.c.7	In-service training (12m)	SEA-PLM, PISA, ERCE, PASEC, PILNA	Partial. Most assessments exceed the 12-month window (covering 2–3 years) or lack detail on the specific types of training.

Note: while PILNA was included in the review of context questionnaires, it is not part of the analysis for this project.

4.2 Questionnaire Frameworks

4.2.1 Conceptual and Theoretical Frameworks in International Large-Scale Assessments

The need for sound and inclusive conceptual and theoretical frameworks in international large-scale assessment (ILSA) studies originates in the early developments of the 1960s, when foundational ideas such as the opportunity to learn (OTL) and the initial input–process–output paradigm first emerged to explain learning outcomes within education systems. The importance of contextual questionnaire frameworks for guiding ILSA design and supporting cross-national comparability is well established in the literature (Kuger, Klieme, Jude, & Kaplan, 2016; Van de Vijver, F. J. R., Jude, & Kuger, 2019).

Conceptual and theoretical frameworks constitute a foundational element of international large-scale assessments (ILSAs), as they define the constructs to be measured, structure the relationships between contextual variables and learning outcomes, and guide the interpretation of results.

4.2.2 Historical Foundations: From OTL to System Models

Carroll's (1963) time-based model posited that achievement reflects the relation between time spent and time needed to learn, laying the groundwork for linking instructional exposure and achievement, providing the core of OTL, and subsequently operationalized in early IEA studies and elaborated in reviews and retrospectives. Building on these foundations, OTL matured as an organizing construct for cross-national comparisons in the IEA's SIMS/TIMSS tradition, emphasizing curriculum content coverage and structured measurement of exposure. It has been repeatedly argued as necessary for valid cross-system inference (Scheerens, 2016; Schmidt & Mcknight, 1995; Wagemaker, 2020c; Wagemaker, 2020a, 2020a).

Historically, the move from early OTL and input–process–output conceptions (Schmidt & Mcknight, 1995) toward dynamic (Creemers & Kyriakides, 2008), multilevel frameworks (Howie, 2022; Kline, 2015) reflects the growing ambition and complexity of ILSAs. Moving from describing what students know to explaining why outcomes vary across contexts, toward how policy and practice might improve them (Plomp, 1998; Scheerens, 2016; Wagemaker, 2020a). This trajectory became critical to provide high quality research characterized by reliable and valid measures and theory grounded interpretations. In summary, the origin of ILSA frameworks has its roots in classic learning theory and OTL (Schmidt & Mcknight, 1995). However, their necessity stems from the demands of valid cross system measurement (Wagemaker, 2020a), multilevel causal reasoning (Kline, 2015), and policy relevant, theory based interpretation that avoids simplistic comparisons (Scheerens, 2016).

Therefore, as ILSAs have grown in prominence and influence globally, these early models have evolved into more complex and comprehensive theoretical structures, incorporating dynamic and reciprocal relationships informed by economic and systems theories. Frameworks such as the dynamic model of school learning developed by Creemers and Kyriakides (Creemers & Kyriakides, 2008) exemplify this progression and have been applied to varying extents across major international studies. The increasing sophistication of ILSA design, anchored in a shared understanding of educational structures and processes across diverse regional and national contexts, underscores why robust theoretical foundations are essential.

In the following sections, the historical development of the IEA studies frameworks is outlined in section 4.2.3.1, followed by the presentation of the current TIMSS and PIRLS frameworks (4.2.3.4, 4.2.3.5).

4.2.3 IEA TIMSS & PIRLS Questionnaire Frameworks

The early IEA research adopted the input–process–output model, which structured educational systems into components such as student background (inputs), instructional processes (teaching and learning), and achievement outcomes. This model provided a first systematic attempt to explain differences in performance across countries by linking outcomes to both system-level and classroom-level factors.

Over time, the OTL concept became central to cross-national comparisons, particularly within the IEA tradition (e.g., SIMS, TIMSS). It emphasized the importance of curriculum coverage and instructional exposure as prerequisites for valid comparisons of student achievement across systems. Without accounting for differences in what students have had the opportunity to learn, cross-country comparisons risk misinterpretation.

4.2.3.1 *Historical Development and Purpose*

Over the past 60 years, IEA has applied large-scale research methods through (with a few exceptions) a curriculum-based, explanatory design, which has varied in form and content across studies (Blömeke, S., Nilsen, T., Olsen, R.V., Gustafsson, J.E., 2022).

A key innovation of IEA frameworks is the conceptualization of curriculum as a multi-level construct. This is most prominently reflected in the distinction between (Blömeke et al., 2022):

- Intended curriculum (system level): national goals, standards, and curricular prescriptions
- Implemented curriculum (school/classroom level): what is actually taught by teachers
- Attained curriculum (student level): what students have learned, including knowledge, skills, and attitudes

This tri-partite model, first formalized in the context of SIMS, has remained a defining feature of IEA studies. It provides a structured way to analyze discrepancies between educational intentions, instructional practices, and learning outcomes, and to compare these relationships across countries.

In addition to curricular content (what is intended to be taught, what is actually implemented in schools and classrooms, and what students ultimately experience and learn), IEA frameworks also incorporate:

- Curricular antecedents (background characteristics, prior conditions, and enabling factors at the system, school, classroom, and student levels that precede educational processes), such as socioeconomic background, home environment, and system-level characteristics
- Curricular contexts (the social environments, organizational arrangements, practices, and behaviors through which teaching and learning take place), including school resources, classroom processes, and student behaviors

Together, these elements form a comprehensive system model that allows for relational analyses linking background factors, educational processes, and outcomes. Importantly, this framework acknowledges that learning is shaped by multiple interacting influences operating at different levels of the education system.

Such a design also enables the development and refinement of predictive models, which utilize data collected at teacher and student level into one conceptual model that includes background influences on schooling, school processes, and student outcomes. Although each individual IEA study develops its own conceptual framework, these typically do not include specifications of indicators but remain on a more general level directed by the IEA’s curriculum-driven underpinning (Howie, 2022).

To conduct relational analyses, it is necessary to specify explicitly what factors are theorized to influence educational achievement either directly or indirectly (Howie, 2022).

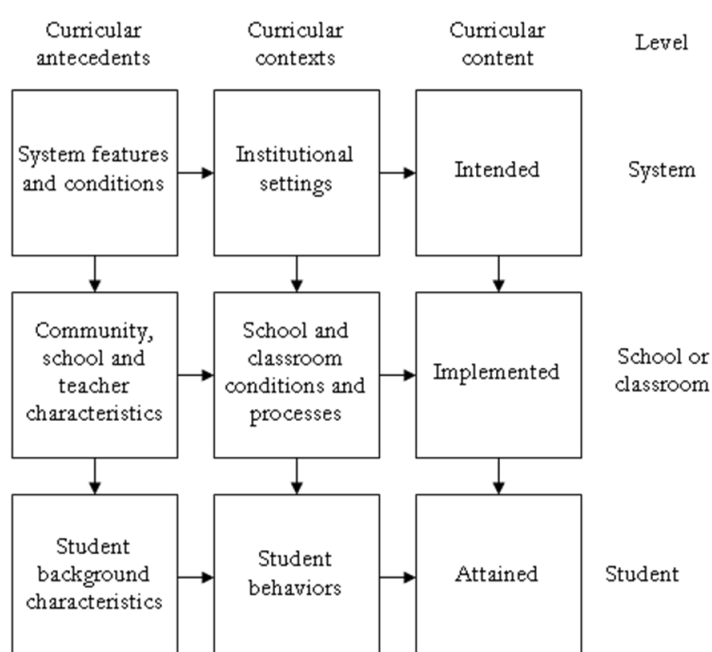


Figure 4-1: Model of an IEA Research Study (IEA 1998, p. 53; Travers and Westbury, 1989 in Plomp et al., 2003).

While this model presented in Figure 4-1 was presented only in 1989, nearly thirty years after the first IEA studies started, it evolved from the central concept of the curriculum and its important role in IEA studies. This is evident from FIMS (1959–1967), FISS (1966–1973), SIMS (1976–1987), and SISS (1980–1989) and was also reflected in TIMSS since its inception in 1995. The IEA regarded the concept of curriculum as being at the heart of its conceptualization of any investigation and evaluation of the teaching and learning of subject matter content at school level (Hastedt, D., Sibberns, H. [2022], 2022). In Travers & Westbury’s model, the curriculum is presented on three levels: firstly, as authorized at the system level (i.e., the intended curriculum), the curriculum as taught by teachers in classrooms (i.e., the implemented curriculum), and the curriculum as learned by the students (i.e., the attained curriculum) (Robitaille & Beaton, 2002).

4.2.3.2 Increasing Complexity: From Static to Dynamic and Multilevel Models

As ILSAs evolved, their frameworks expanded in scope and complexity. The focus shifted from descriptive comparisons of achievement (“what students know”) to explanatory and policy-relevant analyses (“why outcomes differ” and “how systems can improve”).

This shift is reflected in the development of:

- Multilevel frameworks, which explicitly model the hierarchical structure of education systems (students nested within classrooms, schools, and systems)
- Dynamic models, such as the model of Creemers and Kyriakides (2008), which conceptualize educational processes as evolving and reciprocal rather than linear

These developments align ILSA frameworks more closely with educational effectiveness research (EER), emphasizing causal pathways and interactions between factors. They also support more sophisticated analytical approaches, such as multilevel modeling, which are now standard in ILSA data analysis.

Overall, the evolution from early OTL and input–process–output models to dynamic, multilevel frameworks reflects both the increasing ambition of ILSAs and the growing demand for policy-relevant insights.

4.2.3.3 TIMSS

Early IEA international large-scale assessments (ILSAs) laid the foundation for the studies conducted from the late 1990s onward. Among these, the first cycle of the Trends in International Mathematics and Science Study (TIMSS) in 1995 was particularly influential. TIMSS 1995 represented a major methodological advance, implemented on a large scale across 45 education systems and employing rigorous standardized procedures, a rotated assessment design, and sophisticated analytical methods.

4.2.3.4 TIMSS Framework: Continuity and Adaptation

The TIMSS framework exemplifies the continuity of the IEA tradition while incorporating gradual modernization. At its core, TIMSS retains the curriculum-based structure and the OTL perspective, maintaining the distinction between intended, implemented, and attained curriculum.

In its earlier formulations (e.g., TIMSS 1995), the framework explicitly linked system-level factors (e.g., policies, curricula), school and classroom processes (e.g., teaching practices), and student characteristics (e.g., prior knowledge, attitudes) within a coherent explanatory model.

In more recent cycles (e.g., TIMSS 2023), the framework is presented in a more modular way, distinguishing between assessment frameworks (content and cognitive domains) and contextual frameworks. The contextual framework organizes variables across:

- Education system level (national contexts, system organization),
- School and classroom level (resources, climate, instruction),
- Student and home level (background, attitudes, learning environment).

While the presentation has become more concise, the underlying conceptual logic remains largely unchanged. The framework has been updated to reflect evolving educational realities, including the integration of information and communication technologies (ICT), increased attention to school climate, safety, and bullying, and the inclusion of pandemic-related disruptions.

4.2.3.5 PIRLS Framework: A Literacy-Specific Multilevel Model

The PIRLS framework shares the multilevel structure of TIMSS but is tailored specifically to reading literacy. It conceptualizes reading as a constructive and interactive process that is shaped by multiple contexts and influenced by both cognitive and affective factors.

PIRLS defines reading literacy broadly as the ability to understand and use written language for a variety of purposes, including learning, participation, and enjoyment. This definition explicitly incorporates comprehension processes, reading purposes, and attitudes toward reading.

The PIRLS conceptual framework emphasizes the role of:

- National and community contexts (e.g., language environment, education system),
- School and classroom contexts (e.g., instruction, resources, teacher preparation),
- Home environments (e.g., literacy activities, parental support),
- Student characteristics (e.g., motivation, attitudes).

A distinctive feature of PIRLS is its strong focus on early literacy development and the interaction between home and school environments. The framework also explicitly acknowledges the nested nature of these contexts and their combined influence on reading outcomes.

4.2.4 Regional and LMIC-Focused Frameworks

Regional ILSAs, particularly those focused on low- and middle-income countries (LMICs), have developed somewhat later and often reflect different priorities and contexts. Examples include PASEC, SACMEQ, LLECE/ERCE, and SEA-PLM.

While these frameworks are broadly aligned with the IEA tradition, they tend to place greater emphasis on contextual constraints (e.g., infrastructure, teacher availability), include variables capturing access and participation (e.g., absenteeism, repetition), and address socioeconomic and cultural factors more explicitly.

They also often respond more directly to policy needs, including system monitoring, capacity building, and evaluation of education reforms. Global initiatives such as AMPL aim to improve comparability across different assessments by defining common proficiency standards and linking diverse assessment systems to a shared framework.

4.2.4.1 Functions and Importance of Frameworks

Across all ILSAs, conceptual frameworks serve three key functions:

1. Ensuring construct validity and comparability

Frameworks define the constructs to be measured and their relationships, thereby reducing the risk of construct drift and cultural bias in cross-national comparisons.

2. Enabling explanatory and policy-relevant analysis

By linking achievement outcomes to contextual variables, frameworks support interpretation beyond descriptive rankings and enable more meaningful policy insights.

3. Aligning design and methodology

Frameworks provide the basis for instrument development, sampling design, and analytical approaches, particularly in multilevel modeling contexts.

Without such frameworks, ILSAs would lack the conceptual coherence necessary for rigorous comparative research and would risk producing results that are difficult to interpret or use for policy purposes.

4.2.4.1.1 Implications for the Present Study

The review of frameworks suggests two main implications for the present work. First, the TIMSS and PIRLS frameworks provide a robust and widely accepted structure for organizing questionnaire content and linking contextual variables to learning outcomes. Second, regional and LMIC-focused frameworks offer important extensions by capturing context-specific factors that may be underrepresented in traditional IEA instruments but are highly relevant for interpreting learning conditions in LMIC contexts.

Combining these perspectives allows for the development of a framework that is both internationally grounded and sensitive to the specific conditions of LMIC contexts, thereby enhancing the relevance and applicability of the proposed questionnaire item sets.

4.3 From Frameworks to Questionnaires

While the previous section examined the conceptual and theoretical frameworks underlying the reviewed international and regional large-scale assessments, the present section shifts attention to their operationalization through contextual questionnaires. In international and regional large-scale assessments, contextual information is typically collected through questionnaires administered to different respondent groups, most commonly students, teachers, school principals, and, in some studies, parents or caregivers. These questionnaires constitute the primary instruments through which contextual information is collected and transformed into indicators and scales for subsequent analysis.

Although questionnaire instruments vary considerably across studies in terms of scope, wording, emphasis, and respondent coverage, they broadly reflect the contextual dimensions addressed within the underlying frameworks. Questionnaire development is also shaped by practical, regional, and policy considerations, meaning that not all questionnaire content maps directly or exclusively onto predefined framework domains.

Consequently, the framework analysis and the questionnaire analysis in this study were conducted as related but distinct analytical steps. Whereas the framework review focused on how learning and contextual factors are conceptualized at a theoretical level, the questionnaire analysis examined how these concepts were operationalized through measurable questionnaire items across assessment programs in order to identify recurring indicators and contextual dimensions.

5 Methodology

This chapter describes the data sources, materials, and analytical procedures underpinning the study. It first describes the types of documentation and data available for each international and regional assessment included, as these defined the scope and level of analysis that could be applied. The analytical approach adopted in this study was dependent on the availability, level of detail, and type of documentation and data across the different assessments, including questionnaire instruments, framework documents, codebooks, and data sets.

Building on this foundation, this chapter outlines the procedures used to compare and align the selected regional and international assessment context questionnaires at the construct level, with the aim of identifying common and complementary thematic areas relevant to learning contexts in LMICs. These construct areas were subsequently differentiated into specific *indicators*, which formed the primary unit of analysis. For each indicator, existing questionnaire items from international and regional assessments were reviewed, and one or more candidate items were selected based on predefined criteria related to conceptual fit, simplicity, and expected usability in LMIC contexts (see a detailed description of the selection criteria below). Where suitable items were not directly available, newly developed questions were proposed. Following an internal review process by the contributing authors, who are partly involved in IEA TIMSS and PIRLS and regional questionnaire development, illustrative question wordings were proposed, following established IEA questionnaire design conventions to support clarity and comparability.

This chapter further describes the use of available psychometric information from prior assessment cycles to flag potential measurement limitations, as well as the incorporation of qualitative feedback from the NRCs who were responsible for implementing the LaNA 2023 Linking Study in participating countries. Results from the psychometric and NRC review processes resulted in a risk assessment, intended to inform organizations considering using the suggested indicators and questions to collect contextual information in LMICs. The following section thus provides a transparent account of how indicators and illustrative questionnaire operationalizations were derived, reviewed, and assessed, while acknowledging constraints arising from uneven documentation availability across assessments, such as cases where only metadata or codebook information was accessible and item-level wording could therefore not be directly examined.

5.1 Data Sources and Materials Used

This study draws on a set of regional and international learning assessments conducted at or relevant to the end of primary education, including TIMSS 2023, PIRLS 2021, LaNA 2023 Linking Study, ERCE 2019, PASEC 2019, SEA-PLM 2019, SACMEQ IV, PISA-D, and AMPL. For each assessment, available documentation and data were reviewed for the subsequent framework comparison, indicator identification, and development of question suggestions, as well as item-level and empirical analyses. While earlier TIMSS and PIRLS cycles may contain additional items relevant to LMIC contexts, the review focused on the most recent accessible materials from each program to maintain a manageable and consistent evidence base. This approach also ensured that more recent areas of emphasis, including teaching processes and learning environments, were adequately represented.

Across assessments, the primary materials retrieved and analyzed included questionnaire framework documents, student, teacher, and school questionnaires, accompanying codebooks

or metadata, and anonymized data sets. The analytical focus of the questionnaire review was the student, teacher, and school questionnaires. Information from parent questionnaires was considered selectively to assess the coverage of specific indicators across assessments; however, parent-level instruments were not a primary focus of the analysis.

For most of the analyzed assessments—including TIMSS 2019, PIRLS 2021, ERCE 2019, PASEC 2019, SEA-PLM 2019, and PISA-D—comprehensive documentation and microdata were available, enabling framework-level, item-level, and empirical analyses. For the LaNA 2023 Linking Study, no standalone questionnaire framework document was available; instead, the LaNA questionnaires are explicitly grounded in the established TIMSS and PIRLS context questionnaire frameworks. For SACMEQ IV, a dedicated questionnaire framework document was not available, and framework information was derived from existing documentation.

For AMPL, codebook information and data were available from the AMPL-ab study conducted in mid-2023 (UNESCO, 2024), but because there was no update to AMPL’s predecessor MILO study framework (UNESCO, UNICEF, World Bank, OECD: GPE, & ACER, 2021), and AMPL instruments had since been adapted, the corresponding framework has not been included in this analysis. For SACMEQ IV, codebook information was available from 2013 (SACMEQ, 2023), but detailed questionnaire instruments for AMPL and SACMEQ were confidential and therefore not accessible to the project. This limited the use of these assessments for the selection process of suitable questions for certain indicators.

Table 5-1 summarizes the availability of frameworks, questionnaires, datasets, and supporting documentation across the reviewed assessments.

Table 5-1: Availability of Frameworks, Questionnaires, Data, and Documentation

Study	Framework Documents	Student Questionnaires	Parent Questionnaires	Teacher Questionnaires	School Questionnaires	Codebook / Metadata	Data Sets	Comment
TIMSS 2019	✓	✓	✓	✓	✓	✓	✓	
PIRLS 2021	✓	✓	✓	✓	✓	✓	✓	
LaNA 2023 Linking Study	–	✓	–	✓	✓	✓	✓	Framework based on TIMSS/PIRLS Framework
ERCE 2019	✓	✓	✓	✓	✓	✓	✓	

Study	Framework Documents	Student Questionnaires	Parent Questionnaires	Teacher Questionnaires	School Questionnaires	Codebook / Metadata	Data Sets	Comment
PASEC 2019	✓	✓	✓	✓	✓	✓	✓	
SEA-PLM 2019	✓	✓	✓	✓	✓	✓	✓	
SACMEQ IV	-	✓	✓	✓	✓	X	✓	
PISA-D	✓	✓	-	✓	✓	✓	✓	
AMPL	/	✓	-	✓	✓	X	✓	

✓: available and used

–: not administered

X: not available to this research project

/: excluded from the analysis

5.2 Framework Comparison and Analytical Approach

In this section, the approach to analyze and compare the international and contextual frameworks is described.

The IEA 1989 framework (Travers, K. J., & Westbury, I., 1989) was used and provided the analytical framework for comparison. The choice of the IEA generic framework as the analytical grid was due to the holistic nature of the original framework that underpinned the IEA studies historically. Secondly, it provided a systematic framework to evaluate the existing IEA studies' frameworks being included in this project and realizing that almost all the studies included in framework analysis had been informed directly or indirectly conceptually by the IEA. Thirdly, the flexibility of the domains and constructs reflected in the framework had historically served the purpose of serving the needs of the continuum of stakeholders in education, including researchers and policymakers.

The IEA's systemic 3-level model (as described in chapter 4.2.3) differentiates between curricular antecedents, curricular context and curricular contents. It provided a comprehensive basis for reviewing and analyzing the international and regional frameworks. The framework comparison set out to analyze the following frameworks: AMPL, ERCE, PASEC, PISA-D, PIRLS, SACMEQ, SEA-PLM and TIMSS.

The analytical approach to analyze and compare the frameworks is described in the sub-sections to follow. Firstly, those frameworks that were available and could be analyzed as well as the approach taken to those without a framework are described. Then the level of analysis that was undertaken is outlined. The process of mapping, harmonization, clustering the frameworks, and finally the identification of the commonalities and differences across the projects frameworks is described.

5.2.1 The Analyzed Frameworks

As described in section 5.2, a variety of different sources from each study were used to analyze the different studies' frameworks. Table 5-2 shows presents the sources of information used for the analysis of contextual frameworks. It the use of/availability of theoretical or conceptual frameworks describing the overall study conceptualization; contextual frameworks describing questionnaire structure and content; and international and technical reports, which in some studies also contain information related to the framework descriptions. The following projects were analyzed in terms of their conceptualization of their contextual data: ERCE, PASEC, PISA-D, PIRLS, SACMEQ, SEA-PLM and TIMSS. Although UIS has designated the COVID-19 MILO Contextual Framework (UNESCO et al., 2021) as AMPL's official framework, the current AMPL-ab questionnaire no longer includes COVID-specific items. This creates a mismatch between the framework and the instrument. For this reason, AMPL was excluded from the framework comparison.

Since no explicit contextual/conceptual frameworks of SACMEQ could be located, the thematic analysis reconstructed an outline of a framework based on the international and technical reports available. Variables were categorized through a thematic content analysis according to the IEA framework. The outline of SEA-PLM's contextual framework was available and underpinned by recent literature although at the time of writing the contextual framework had explicit thematic areas, but the constructs and variables were suggestive of what was to be measured. While ERCE's schematic framework was available and some aspects were elaborated, the empirical underpinning based on the literature was less evident with the factors and variables being only exemplars of what was to be measured. PASEC also provided an explicit contextual/theoretical framework described in the PASEC 2019 Technical Report, structured across student/family, school/classroom, and system levels and organized around inputs, processes, perceptions, and outputs. The framework was designed to reflect the educational realities and equity challenges of Francophone Sub-Saharan Africa and provided a clear conceptual linkage between contextual variables and learning outcomes. The contextual frameworks for PIRLS, PISA-D and TIMSS were located and analyzed and supplemented by international and technical reports. PIRLS and TIMSS included domains and constructs

Table 5-2: Source of Information Used for Thematic Analysis of Projects' Contextual Framework/Data

Study	Theoretical/ Conceptual Framework	Contextual Framework	International / Regional Report	Technical Report	Comment
TIMSS 2023	✓	✓	✓	✓	Detailed contextual framework formed basis for analysis and inclusion.
PIRLS 2021	✓	✓	✓	✓	Detailed contextual framework formed basis for analysis and inclusion.
ERCE 2019	✓	✓	✓	✓	Detailed contextual framework formed basis for analysis and inclusion supported by information from regional report
PASEC 2019	✓	✓	✓	✓	Detailed contextual framework formed basis for analysis and inclusion.
SEA- PLM 2019	n/a	✓	✓	✓	Contextual framework formed basis for analysis and inclusion supplemented by regional report
SACMEQ IV	-	-	✓	✓	Topics/sub-topics reconstructed from regional and technical reports. No conceptual/contextual framework could be found.
PISA-D	n/a	✓	✓	✓	Contextual framework formed basis for analysis and inclusion supplemented by international and technical report.
AMPL	-	-	-	✓	No actual contextual framework: mismatch between official framework (COVID-19 MILO) and revised questionnaire (COVID items removed); reflects instrument still under development

(✓): available and used

(-): Not found

(n/a): Not applicable/not used

5.2.2 Level of Comparison (Domains, Constructs, Thematic Areas)

The IEA 1989 generic framework (see description in section 4.2) served as an analytical framework and provided the basis for the comparison of the various frameworks that were captured in Excel (see Table 5-3). The analytical grid evaluated the contents of the frameworks in terms of the systemic level (System, School and Student), the curricular content, the curricular antecedents and curricular contexts as far as could be identified and interpreted.

Each of the frameworks was located and analyzed in terms of their thematic areas. Where contextual frameworks were found, the framework was analyzed in terms of the broad areas and whether these were differentiated systemically. Most of the explicit frameworks contained a systemic approach with multiple levels (e.g., National, School and/or Class and Home/individual level), with the exception of the PISA-D framework. For ERCE, PASEC, PISA-D, PIRLS, SEA-PLM, and TIMSS, each framework's themes/thematic areas, topics and sub-topics were reviewed, categorized and aligned with the IEA framework and with each other wherever feasible.

Table 5-3: Analytical Grid Used for Analysis and Comparison of Frameworks

Level	Curricular Content	Curricular antecedents	Curricular contexts	TIMSS (2023)	PIRLS 2021	AMPL (Assessment of Minimum Proficiency Levels)	SEA-PLM Southeast Asia Primary Learning Metrics (2019)	SEACMEQ IV The Southern and Eastern Africa Consortium for Monitoring Educational Quality	PASEC Programme for the Analysis of Education Systems 2019	ERCE Regional Comparative and Explanatory Study (2019)	PISAD (Program for International Student Assessment - Development)
System	<u>Intended</u>										
		System Features and conditions									
		Institutional Settings									
School or Classroom	<u>Implemented</u>										
		Community, School and teacher characteristics									
		School and Classroom conditions and processes									
Student	<u>Attained</u>										
		Student Background characteristics									
		Student behaviors									

5.2.3 Alignment of Terminology (Mapping, Harmonization, Clustering) and LMIC-Focused Frameworks

The nomenclature and terminology varied across the projects. Using the IEA framework as the basis for categorization, the alignment considered the differences in concepts and the wording thereof between projects. For example, the student level conceptualization and associated terminology had to distinguish between home-focused environments as opposed to the student's own attributes or characteristics. Students (IEA framework terminology) were referred to as learners or pupils or "individuals" ranging from student attributes, antecedents: student characteristics, pupil characteristics, student level demographic factors but were all categorized under the student-level category and Student behaviors in the IEA framework which had specified student-specific characteristics as opposed to the home environment. Similarly, Home environment related themes and topics across projects varied in reference to parents, home, home contexts, home environment/family background, families. These were all categorized under Student-level and Student background characteristics.

All projects' frameworks or reconstructed outlines in place of a framework were mapped onto the IEA 1989 framework (see section 6.2).

In the analytical grid designed based on the IEA framework, all thematic areas/themes, topics were clustered within the IEA categories and levels. Harmonization was mostly achieved by re-aligning each of the project's themes/topics in rows thematically. For example, Student attributes within which Student Demographics was a sub-topic included in several frameworks. Student Demographics appeared in both TIMSS and PIRLS and were aligned under Student level. TIMSS in its framework referring to age, gender, country of birth within the text of the framework in contrast to PIRLS which under Student Demographics elevated and listed gender and age as sub-topics, both were aligned in parallel with each other. Other projects with student demographics as a topic (SEA-PLM, PASEC, PISA-D) were aligned similarly. Those projects where the frameworks were implicit, but reports referred to Pupil (student in other frameworks) characteristics (SACMEQ) and where variables of gender, age or other personal characteristics related specifically to the students tested, *these were clustered and harmonized in the framework under the same student level domain.*

5.2.4 Identification of Common vs Complementary Themes

By mapping, aligning and harmonizing the thematic areas and topics onto a common framework systemically, common themes as well as differences in emphases between projects became evident, and these are reported in section 6.2. A systematic thematic analysis and comparison of the thematic areas/topics within projects with formal frameworks or clustered and categorized mapped variables from projects without explicit frameworks, provided evidence of the different emphases and approaches to the international and regional studies.

5.3 Conceptual Review, Indicator Construction, and Item Selection

On questionnaire level, the conceptual phase of the study focused on deriving indicators and identifying and refining the corresponding questionnaire items prior to any empirical analysis. The objective was to ensure 1) relevance for LMIC contexts, 2) continuity with established IEA assessment frameworks, and 3) alignment with SDG reporting priorities and stakeholder requests.

Indicator development followed a structured, multi-stage approach that combined systematic framework comparison, database analysis, and a transparent scoring mechanism. Candidate

indicators were assessed primarily according to two guiding criteria: their demonstrated or plausible association with student learning outcomes and their relevance for education policy and monitoring in LMIC contexts. In this way, priority was given to constructs that have been identified in the literature and in ILSAs and regional assessments as influencing learning conditions, while also reflecting areas of policy interest such as school resources, instructional practices, teacher capacity, and student learning environments.

Evidence from the literature review (see Chapter 4) and from existing ILSAs and regional assessments informed the identification and interpretation of relevant constructs. Findings from these assessments and the related research were used to identify factors that have been shown to be associated with learning outcomes, while experience from questionnaire development in IEA studies and regional assessments helped ensure that selected indicators could be operationalized in a feasible and meaningful way within context questionnaires. Professional judgment of the project team—drawing on extensive experience in IEA questionnaire development and assessment work in LMIC contexts—was therefore applied throughout the process to ensure that strategically important constructs were not overlooked where empirical coverage across sources was limited.

For each respondent level (student, teacher, and school), questionnaire content from selected regional and international assessments was systematically reviewed. Items were grouped into conceptually coherent indicators that aligned with the subdomains identified during the framework comparison process. Following the construction of these indicators, questionnaire items corresponding to each indicator were identified, compared across questionnaire sources, and evaluated with respect to clarity, feasibility, and alignment with established IEA terminology. Where necessary, wording adjustments were proposed to improve contextual relevance and harmonization between items while maintaining conceptual comparability.

This process resulted in a structured shortlist of indicators and a corresponding pool of candidate questionnaire items, which formed the basis for the subsequent selection steps.

5.3.1 Database Selection and Scoring Strategy

The conceptual review began with the systematic identification and retrieval of context questionnaire instruments, technical documentation, and, when possible, publicly available codebooks from ILSAs and regional assessments. Data and documentation were obtained from official study websites. Where required, permission to use materials for this research project was obtained from the respective data custodians.

The review focused on assessments measuring mathematics and/or reading at or near the end of primary education. Selection criteria for source databases were designed to maximize coverage of constructs relevant to primary-level learning environments in LMICs while maintaining conceptual alignment with established IEA frameworks.

The purpose was to identify constructs, indicators, and corresponding questionnaire items reflecting policy priorities, feasibility in LMIC contexts, and continuity with international monitoring efforts. Consultations and technical discussions with the World Bank and UNESCO further informed the identification of policy-relevant domains and SDG-linked constructs.

5.3.1.1 Assessments and Initiatives Included in the Selection Process

Indicators and candidate items were extracted from the following sets of instruments and initiatives:

1. **Regional LMIC assessments and initiatives:**

SEA-PLM, PASEC, ERCE, and SACMEQ were each included as source review databases. These assessments were selected because they directly reflect educational conditions and policy priorities in LMICs, offering repeated, context-specific evidence of feasibility and relevance. Together, they cover major geographic regions (Sub-Saharan Africa, Latin America, and Southeast Asia) and have produced stable and well-documented context questionnaire frameworks.

The Pacific Islands Literacy and Numeracy Assessment (PILNA) was not included in the current review due to its comparatively small and highly specific population base, which limits the generalizability of its contextual measures to the broader LMIC cohort targeted by this study.

2. **IEA mathematics and reading assessments:**

TIMSS and PIRLS were included to ensure continuity with widely used international frameworks. LaNA was included as the IEA assessment specifically designed for LMIC contexts, serving as an important reference point for indicator alignment.

3. **International LMIC-focused initiatives:**

AMPL and PISA-D were included to capture additional LMIC-specific measurement approaches and policy priorities. Although AMPL provided a more limited set of contextual constructs, it was explicitly developed for primary-level LMIC contexts and incorporated elements aligned with selected SDG indicators. For this reason, it was included as a reference source in the conceptual review.

PISA-D, while targeting an older student cohort, was primarily reviewed for constructs and item formulations at the teacher and school levels, where age-related differences are less consequential. Its inclusion acknowledged its bridging role between LMIC realities and global policy frameworks and its contribution to policy-relevant contextual measurement.

5.3.1.2 Indicator Generation

For each analytical level (student, teacher, and school), all questionnaire items extracted from the respective instruments (or from codebooks in the case of AMPL and SACMEQ) were grouped into distinct indicators. These indicators were aligned, as far as possible, with the subdomains identified during the framework review and comparison process. For this project, home/parent questionnaire data were not treated as a separate level but were integrated at the student level, reflecting study design and data merging conventions.

The indicator construction process drew initially on the TIMSS 2019 context questionnaires, which provided the comparative reference structure for general background and mathematics-related constructs. PIRLS 2021 contributed complementary reading-specific constructs. This initial set of indicators was subsequently extended through the integration of relevant constructs from the LaNA student and school context questionnaires to strengthen contextual relevance and applicability for LMIC settings.

The resulting consolidated indicator set served as the reference framework against which all regional questionnaire items were mapped. During an intensive review process, the authors sought to balance conceptual granularity with cross-study comparability. This involved ensuring sufficient differentiation between indicators while maintaining coherent alignment across instruments and preserving consistency with established IEA terminology where appropriate.

The procedure was iterative. Some indicators were disaggregated to better reflect distinct constructs emerging across instruments, while others were merged to avoid conceptual redundancy and excessive fragmentation. In several cases, indicator labels were refined to ensure that they adequately captured the full semantic scope of items originating from different assessments and to enhance conceptual clarity.

5.3.1.3 Indicator Scoring, Inclusion and Exclusion Criteria

Once a complete mapping between the defined indicator set and corresponding regional and international questionnaire items had been established, additional priorities were integrated. Specifically, SDG indicators measurable through ILSA background instruments (see Section 4.1.4) were linked to the relevant indicators to ensure alignment with global monitoring frameworks and to enhance policy relevance, particularly for LMIC contexts. In parallel, partner requests, especially those from the World Bank and UNESCO, were incorporated and explicitly flagged to reflect their strategic importance for global monitoring and policy dialogue.

5.3.1.3.1 Scoring

One point was awarded for each regional assessment contributing at least one matching item (or response category), including AMPL and PISA-D. TIMSS and PIRLS were treated jointly for scoring purposes: a maximum of one additional point was assigned across these two studies, reflecting the substantial overlap in general background constructs between the IEA instruments. These differences between TIMSS and PIRLS are largely an artifact of cycle-specific instrument development following the 2011 harmonization of the two studies.

Applying the above scoring criteria, an indicator at the student and school levels could achieve a maximum of nine points: up to seven points for contributions from the different assessment programs, one point for SDG relevance, and one point for alignment with partner requests.

As AMPL does not include a teacher questionnaire, indicators at the teacher level could achieve a maximum of eight points.

5.3.1.3.2 Indicator Selection

Indicators were included in a shortlist for further review if they met one or both of the following criteria:

1. The indicator achieved a score above the predefined threshold: more than three points (out of a maximum of nine) at the student and school levels, or more than two points (out of a maximum of eight) at the teacher level. Scores above these thresholds indicate representation across multiple independent regional or international assessments and/or strong policy relevance through alignment with SDG targets and/or partner-identified priority areas.
2. Certain indicators that did not meet criteria 1 or 2 were nevertheless retained based on the professional judgment of the author team. This override mechanism was applied throughout the review process to preserve indicators considered strategically important, such as those addressing equity, emerging policy priorities, or recurrent issues in regional assessments targeting LMICs, and also included established indicators already included in the IEA LaNA 2023 Link Questionnaires, which were specifically designed for the LMIC context.

As a result, an indicator shortlist was developed, comprising a structured pool of indicators at the student, teacher, and school levels. The shortlist was reviewed and discussed within the author team, and following several rounds of refinement, it was consolidated into a structured indicator database. This database served as the basis for identifying and selecting specific

questionnaire items from relevant assessment instruments corresponding to each indicator. These items were then considered for recommendation and further development according to the criteria specified in the following section.

5.4 Item Identification, Selection, and Adaptation

Following the identification and selection of candidate indicators, the next stage focused on identifying and refining questionnaire items (i.e., questions) suitable for measuring these indicators in LMIC contexts. This stage involved three main steps: 1) compiling existing items from relevant assessments, 2) selecting those that best represented the defined indicators, and 3) revising or adapting selected items where necessary.

As a first step, potentially relevant items were systematically compiled from the source assessments included in the review (i.e., TIMSS, PIRLS, LaNA, AMPL, SEA-PLM, PASEC, ERCE, SACMEQ, and—only in case other regional initiatives could not provide the necessary information—PISA-D). For each indicator, items that appeared to capture the corresponding construct were documented in a structured item list (Questionnaire_Themes_Overview.xlsx) to support comparison across instruments. The catalogue recorded the assessment source, the respondent group (student, teacher, school questionnaire, or parent where applicable), and the location or item number within the respective instrument. Additional fields allowed members of the research and development team to add comments and note potential preferences or suggestions for rephrasing.

The purpose of this catalogue was not only to compile candidate items, but also to allow a systematic comparison of how similar constructs were operationalized across assessments. This helped identify items that most directly captured the intended indicator while also highlighting cases in which regional assessments measured similar constructs differently from IEA instruments, particularly in domains such as home learning resources, basic school infrastructure, attendance patterns, and language exposure.

In a second step, items from the catalogue were reviewed and selected based on their suitability for measuring the defined indicators. The following criteria guided the selection:

- 1. Leverage existing IEA items first.**

Where a suitable item already existed within TIMSS, PIRLS, or LaNA, it was preferred for comparability reasons to maintain linkage with existing trend data. This approach supports continuity with established IEA constructs and maximizes the potential for cross-country and cross-study comparability. Adaptations were considered when there was evidence that an item had proven difficult to administer, was poorly understood, or showed insufficient discrimination in LMIC contexts, including cases where feedback or psychometric analyses from PIRLS 2021 or TIMSS 2023 informed revisions for PIRLS 2026 or TIMSS 2027. For constructs where regional assessments have established validated operationalizations in LMIC contexts (particularly SES-related, infrastructure, and teacher labor constructs), items from regional sources were selected.

- 2. Choose the simplest effective version across studies.**

Where multiple versions existed across sources, the item with the least complexity was prioritized for selection. For example, when measuring household composition, simpler checklist or count-based questions used in some regional assessments were preferred over full household roster approaches, provided they preserved the essential

information content while also being simple enough to be reliably answered by students at the end of primary education.

3. Target simple and familiar response formats.

Items were designed to be easy to administer and respond to in large-scale settings. Yes/No or multiple-choice items were selected where possible, and agreement or frequency scales were limited to three or four categories. Open-ended formats were avoided, with limited exceptions for low-burden numeric responses that children could answer reliably (e.g., age or simple counts).

4. Avoid complex filters or skip logic (especially for students).

For student instruments, items requiring routing were excluded due to the higher likelihood of implementation and response errors. Teacher and school instruments could include minimal routing, if necessary, but this was treated as a design exception rather than the default. Wherever possible, conditions were incorporated directly into the response options.

5. Minimize cognitive and memory load through design.

Items were phrased to reduce recall burden and abstraction. Where approximate reporting was sufficient, ranges were preferred over exact numbers (e.g., book categories rather than precise counts).

6. Use universally understood concepts and adapt examples to local realities.

Preference was given to concepts likely to translate across contexts (e.g., mobile phone ownership rather than landline access). Where examples were needed, they were designed to be locally adaptable (housing materials, transport modes, food and nutrition indicators) to avoid bias and increase respondent comprehension.

7. Allow regional complementarity to capture LMIC-relevant aspects.

Where the IEA base item provided a strong comparability anchor but did not fully capture LMIC-relevant variation, items from regional assessments served as the primary operationalization. This approach supports both international comparability and contextual sensitivity (e.g., maintaining a core language-of-test item while supplementing questions on multilingual home environments where relevant).

8. Standardize categories where feasible.

To strengthen analytic comparability, the review prioritized alignment with international classifications (e.g., ISCED for parental education) and harmonized category structures across items when comparable information was available across assessments.

In the final step, the selected items were reviewed and, where necessary, revised or adapted. At this stage, items were refined based on the professional judgment of the research team and informed by ongoing discussions within the PIRLS 2026 and TIMSS 2027 Questionnaire Item Review Committee (QUIRC). In some cases, items were rephrased to improve clarity or contextual relevance; in others, additional items were included to ensure that the indicator was adequately covered. The resulting set of candidate items therefore retained maximum comparability with existing IEA instruments while allowing targeted adaptation for LMIC contexts and the LaNA study design (e.g., paper-based administration and varying familiarity with survey response formats).

Given the granular nature of this work, several meetings were convened among the experts to go over the assembled catalogue item by item.

Constructs that were only measured via parent questionnaires in the source assessments were not represented by items on the student level student; rather, these constructs were either excluded from student-level proposals or earmarked for consideration at other respondent levels, depending on feasibility within the broader instrument design.

5.5 Empirical Screening on Question and Indicator Level

Following the completion of the content-based indicator construction and item selection processes described in Section 5.3, an empirical screening was conducted to assess the analytical suitability of the selected variables. This verification step was applied prior to final recommendations and focused on evaluating response quality, distributional characteristics, and basic empirical relevance within LMIC samples.

Table 5-4: Study-Specific LMIC Data used for Empirical Screening

TIMSS	PIRLS	LANA	AMPL	ERCE	PASEC	SACMEQ	SEAPLM	PISA-D
Albania	Albania	Burkina Faso	Kenya	Argentina	Benin	Botswana	Cambodia	Cambodia
Armenia	Azerbaijan	Egypt	Lesotho	Brazil	Burkina Faso	Kenya	Lao PDR	Ecuador
Azerbaijan	Brazil	Nigeria	Zambia	Colombia	Burundi	Lesotho	Myanmar	Guatemala
Bosnia & Herz.	Egypt	Pakistan	Gambia	Cuba	Cameroon	Malawi	Malaysia	Honduras
Brazil	Georgia	Palestine		Dominican Rep.	Rep. of Congo	Mauritius	Philippines	Paraguay
Georgia	Iran	Senegal		Ecuador	Côte d'Ivoire	Namibia	Viet Nam	Senegal
Iran	Jordan			Guatemala	Gabon	Seychelles		Zambia
Jordan	Kazakhstan			Honduras	Guinea	South Africa		
Kazakhstan	Morocco			Mexico	Madagascar	Eswatini		
Morocco	N. Macedonia			Nicaragua	Niger	Uganda		
N. Macedonia	Montenegro			Peru	DR Congo	Zambia		
Montenegro	Serbia			Paraguay	Senegal	Zanzibar		
Serbia	Turkey			El Salvador	Chad	Zimbabwe		
Turkey	Uzbekistan				Togo			
Uzbekistan	Kosovo							
Kosovo	South Africa							
South Africa								

Table 5-5: Number of Shared Countries Across Studies

Shared countries across studies (diagonal = total per study)

TIMSS	17	15	0	0	1	0	1	0	0
PIRLS	15	16	1	0	1	0	1	0	0
LANA	0	1	6	0	0	2	0	0	1
AMPL	0	0	0	4	0	0	3	0	1
ERCE	1	1	0	0	13	0	0	0	4
PASEC	0	0	2	0	0	13	0	0	1
SACMEQ	1	1	0	3	0	0	13	0	1
SEAPLM	0	0	0	0	0	0	0	6	1
PISA-D	0	0	1	1	4	1	1	1	7
	TIMSS	PIRLS	LANA	AMPL	ERCE	PASEC	SACMEQ	SEAPLM	PISA-D

Prior to analysis, countries not classified as low- or lower-middle-income according to the World Bank income classification were removed from the analytical datasets to ensure alignment with the study’s intended policy context. Countries included in the analysis are listed in Table 5-4, while Table 5-5 summarizes the overlap in participating countries across the different assessment programs. As expected, the strongest overlap occurs between TIMSS and PIRLS, with 15 countries participating in both studies. Three of the four countries participating in AMPL also participated in SACMEQ, while four of the seven PISA-D countries also participated in ERCE. In contrast, there is no overlap among the regional assessment programs (ERCE, PASEC, SACMEQ, and SEA-PLM), reflecting their distinct geographical coverage. Only a small number of countries participating in these regional assessments also took part in the international IEA studies.

5.5.1 Scope of Verification and Data Preparation

The empirical verification was conducted separately for each assessment’s dataset and restricted to participating LMIC countries. For each dataset, study-specific student weight variables were applied in all analyses and student achievement was operationalized using study-specific plausible value variables or single-score achievement measures where plausible values were not available

All selected background variables were examined with respect to three core dimensions:

1. Response quality, assessed through item-level missingness patterns
2. Distributional properties, examined through the percentage distribution across response categories for categorical variables
3. Associations with student achievement, used as descriptive indicators of empirical relevance.

Response quality was operationalized at the item level as the proportion of valid responses relative to omitted, invalid, and system-missing responses, using study-specific missing code definitions. This approach ensured comparability across assessments with differing questionnaire designs and coding schemes.

The purpose of this screening was not to conduct full psychometric validation. Rather, it aimed to identify variables that may be analytically weak, insufficiently informative, or potentially misleading when used in cross-country descriptive or policy-oriented analyses.

5.5.2 Handling of Achievement Outcomes and Data Linkage

Student achievement was analyzed separately for mathematics and reading. Where plausible values were available, achievement scores were computed as the arithmetic mean across all plausible values within a domain. Where only single-score achievement measures were available, these were used directly.

Associations were estimated at the appropriate analytical level depending on the background variable under consideration:

- For student-level background variables, achievement outcomes were analyzed directly at the student level.
- For teacher- and school-level background variables, student achievement data were aggregated to the school level and merged with the corresponding background information, ensuring that associations reflect the appropriate hierarchical linkage.

Associations were first computed separately for mathematics and reading. For reporting purposes, and to reduce complexity, the primary indicator presented in the main analysis is an overall association with achievement, defined as follows:

- If both mathematics and reading associations were available, the overall association was computed as the arithmetic means of the two domain-specific association measures.
- If only one domain was available, the overall association was based solely on the available domain.

5.5.3 Association Measure and Statistical Techniques

Associations between background variables and achievement were quantified using the coefficient of determination (R^2) from weighted linear regression models. For each background variable, a bivariate regression of achievement on the background variable was estimated using analytic (weighted least squares) weights, normalized to have a mean of one within each country. This approach ensures comparability across countries and avoids inflation of model fit due to unequal weighting schemes or large sample sizes.

The reported R^2 values are model-based measures of explained variance derived from weighted regression models including an intercept. They are not squared correlations and should be interpreted as indicators of explanatory strength rather than symmetric measures of association. Consequently, a low bivariate R^2 does not necessarily imply that an indicator is unimportant, as its contribution may change when considered jointly with other variables due to mediation, suppression, or shared variance.

Although bivariate correlations can provide descriptive information, associations with achievement were analyzed exclusively using weighted regression models. This approach was chosen because it allows the consistent application of analytic weights, accommodates different measurement levels of background variables, and provides a comparable measure of explained variance across studies and countries. Squared correlations were therefore not used as primary indicators.

5.5.4 Screening Criteria, Reliability and Validity Considerations

For each retained indicator, all corresponding questions selected to measure on this indicator were included for empirical examination and were evaluated against a set of predefined empirical screening criteria listed in Table 5-6. These thresholds were used to flag items for closer review rather than as automatic exclusion criteria. They identify psychometric properties of items that fall outside the usual range and may warrant further examination or adaptation before future use. However, final questionnaire recommendations may deviate from the initially retained items due to subsequent reformulations or adaptations. Items therefore were classified according to their modification level (see section 5.7.1 for more details) and only results obtained from modification level one and two were considered in the final evaluation.

While several item properties were calculated, response quality (missingness) and dominant category distributions were treated as the primary indicators, as they provide the most relevant information for assessing the practical usability of items and identifying potential adaptation needs. The sparse-category criterion was calculated but was not included in the final evaluation, as low-frequency response categories do not necessarily indicate a problematic item. In many cases, categories can be collapsed for analytical purposes, while it is acknowledged here that infrequent responses may nevertheless need to be monitored since they may affect policy-relevant conditions. Calculated psychometric properties for all investigated items therefore can be obtained from the authors upon request. Information about associations with achievement are additionally reported to provide complementary context.

The decision to base the risk indicator solely on missingness and dominant category information reflects a pragmatic prioritization of criteria most directly affecting the practical usability of items. High non-response rates and highly dominant response categories limit empirical coverage and discriminatory capacity in ways that are difficult to address post hoc. In contrast, sparse categories can often be handled through category aggregation, and weak associations with achievement do not necessarily undermine policy relevance.

Table 5-6: Summary of Screening Criteria and Thresholds for Background Variables

Dimension	Indicator	Threshold	Interpretation	Used for Review
Response quality / response rates	Omitted or invalid responses	> 20% of responses missing	Insufficient empirical coverage	Yes
Distribution (categorical)	Dominant category	≥ 80% in a single category	Limited discrimination	Yes
Distribution (categorical)	Sparse category	< 5% in any category	Potentially unstable estimates	No
Association with achievement	Overall $R^2 < 0.01$	Negligible explanatory value		Reported

Note: The indicator thresholds are used to flag items for closer review but do not constitute an automatic exclusion criterion.

5.6 Qualitative Input from LMIC Representatives

Following the completion of the content-based indicator selection and the preliminary empirical screening described in the previous sections, a structured qualitative review was conducted to incorporate contextual expertise from LMIC representatives. This step aimed to assess the practical suitability, clarity, and contextual relevance of the proposed questionnaire item sets prior to final recommendation.

5.6.1 Design of the Review Instrument

A standardized Excel-based review template was developed for the student, teacher, and school questionnaires. For each proposed item, the template listed:

- The suggested item stem
- The corresponding response options
- Three structured rating dimensions, and
- A free-text comment field

The three rating dimensions were defined as follows:

Clarity: Whether the wording and content of the item are understandable for the intended respondent group in the national context.

Differentiation: Whether the item is expected to meaningfully differentiate among respondents rather than produce highly homogeneous responses.

Policy relevance: Whether the information generated by the item is relevant for national education policy or system monitoring purposes.

Each dimension was rated using a three-point response scale (Clarity: “Yes”, “Possibly unclear”, “Likely unclear” / Differentiation and Policy Relevance: “Yes”, “To some extent”, “No”).

These procedures allowed for structured and comparable but low-burden feedback. The comment field enabled reviewers to suggest minor wording adjustments, identify potential response-category issues, or indicate contextual concerns not captured by the rating scales.

5.6.2 General Questionnaire-Level Review

In addition to the item-level sheets, the template included a dedicated “GENERAL” sheet for overall evaluation of each questionnaire level. This section asked reviewers to:

- Assess the overall coverage of key constructs,
- Indicate whether important indicators or domains were missing,
- Provide brief descriptions of any missing elements, and
- Rate the overall national acceptability of administering the questionnaire in its current form.

The overall coverage and the national acceptability of the suggested questionnaire questions were also rated using a three-point response scale (“high”, “medium”, “low”).

Using this structure allowed reviewers to move beyond individual items and reflect on the coherence and completeness of the suggested indicators and questions.

5.6.3 Implementation

The review templates, together with detailed written instructions explaining the rating dimensions and response expectations, were distributed to all six countries participating in the LaNA 2023 Linking Study. Reviewers were explicitly informed that the underlying constructs and indicators had already been selected based on established international and regional assessment frameworks. The purpose of this review was therefore not to reconsider the overall conceptual framework, but to evaluate whether the proposed items are likely to function appropriately within national LMIC contexts.

The qualitative feedback obtained through this process served two identify items that may require revision due to clarity or differentiation concerns and to document contextual considerations relevant for future adaptation and implementation.

The structured ratings were subsequently transformed into quantitative summary indicators to allow integration with the empirical screening results described in Section 5.5 into an item suitability assessment described in the subsequent section.

5.7 Item Suitability Assessment

To support structured and transparent decision-making in the refinement and further use of the proposed questionnaire item sets, an item suitability assessment was developed. The framework integrates two independent sources of evidence:

1. Expert review feedback from LMIC representatives (Section 5.6), and
2. Psychometric screening based on existing large-scale assessment data (Section 5.5).

The purpose of this procedure is to provide structured guidance to researchers using the proposed questionnaire item sets by identifying questions that may require revision prior to finalization or pilot testing. The approach deliberately prioritizes interpretability, transparency, and reproducibility over statistical complexity.

5.7.1 Psychometric Item Suitability

Psychometric evidence was derived from variable-level statistics corresponding to selected items from existing regional or international assessments administered in LMIC contexts. Screening was limited to structural indicators of item functioning. Correlations with

achievement, calculated using LMIC samples from the respective contributing study, are provided as supplementary information but were not used as risk triggers.

At the variable level, a variable was flagged as exhibiting a psychometric issue if at least one of the following conditions was met:

1. High omission rate: $\geq 20\%$ invalid or omitted responses; or
2. Dominant response pattern: $\geq 80\%$ of responses concentrated in a single category.

These criteria capture potential operational limitations such as comprehension difficulties, sensitivity issues, response bias, or limited discriminatory capacity. Measures of association with achievement were retained descriptively but were not used as risk triggers.

For items consisting of multiple variables (e.g., matrix questions), psychometric evidence of a certain risk was determined by aggregating at the question level. For each question the following aggregations were performed:

- N_{vars} = number of associated variables
- N_{risk} = number of variables flagged as exhibiting an issue
- $\text{PCT}_{\text{risk}} = N_{\text{risk}} / N_{\text{vars}}$

A question was classified as exhibiting psychometric risk when more than 50% of its associated variables were flagged. This majority rule prevents isolated sub-item issues from resulting in over-classification while ensuring that systematic weaknesses are identified.

Since many of the selected questions were adapted on the professional judgment of the research team as described in chapter 5.4. Since the items finally suggested may differ significantly from the originally selected existing international and regional questions, the degree of modification was evaluated and classified in three different categories:

- **Group 1: Psychometric Equivalence Expected**
 - **Criteria:** The wording, construct, and response options are nearly identical.
 - **Implication:** Item parameters (difficulty, discrimination) from original studies like PIRLS/TIMSS/LaNA can likely be used directly for trend analysis.
- **Group 2: Psychometric Sensitivity Likely**
 - **Criteria:** The core construct is identical, but category refinements (e.g., more granular ranges) or minor stem adjustments for inclusivity occur.
 - Addition of new categories (up to 3 new categories)
 - Adding new dimensions to the categories
 - Expansion of the amount of Likert type response categories
 - Change numeric count to a categorical checklist or vice versa
 - **Implication:** Thresholds may shift. While the item is “meaningful” for the same latent variable, new calibration is required to confirm the numerical parameters.

- **Group 3: Significant/NEW (Recalibration Required)**
 - **Criteria:** Significant construct shifts (e.g., from health status to classroom impact), synthesis of multiple regional sources, or introduction of entirely new dimensions.
 - Addition of new categories (more than 3 new categories)
 - A change in scope of the answer categories (e.g., from binary “Yes/No” to frequency scale)
 - **Implication:** Previous psychometric parameters are no longer applicable; full validation is required.

Only items of modification group 1 or 2 were regarded as psychometric analyses. Items classified as belonging to group 3 were not regarded and also were not assessed in the overall item suitability assessment described in chapter 5.7

In addition, descriptive statistics were created for reporting purposes, including the average omission proportion, average dominance proportion, and mean association (r^2) across variables.

Details on the psychometric properties of all variables included in the analyses are available from the authors upon request.

5.7.2 Expert Review-Based Item Suitability

For each retained question, expert feedback was collected at the question level across three dimensions: *clarity*, *differentiation*, and *policy relevance*, using a three-point response scale (“Yes”, “Possibly unclear/To some extent”, “No”) which was converted onto a numerical 0-2 scale for analytic purposes.

To identify measurement-related concerns, only *clarity* and *differentiation* were used in constructing an expert-based risk indicator. *Policy relevance* was excluded from the risk trigger and presented as separate dimension, as low policy relevance reflects substantive prioritization rather than measurement weakness.

For each question, the minimum of the average clarity and differentiation ratings was calculated.

A question was classified as exhibiting expert-based measurement risk when the resulting value was below 1.5.

A threshold of 1.5 ensures that items rated consistently as “Yes” are not flagged, while items receiving mixed or predominantly “Possibly unclear/To some extent” evaluations are identified for further consideration. The weakest-dimension approach prevents concerns in one area from being masked by stronger ratings in another.

5.7.3 Treatment of Newly Developed or Substantially Modified Items

For newly developed items or items substantially revised relative to their source assessments, psychometric comparability could not be assumed. These cases were identified through a dedicated comparability flag.

For such items, psychometric risk was not computed, and overall item suitability was marked as not applicable. These items require empirical validation in subsequent pilot testing phases.

5.7.4 Overall Item Suitability

For items with available psychometric evidence (Modification level (MOD) < 3), overall suitability was determined through a combination of expert judgment and psychometric screening results. Items were classified into the following categories:

- **Suitable:** No concerns identified in either expert review or psychometric screening
- **Consider Revision:** Concerns identified in either expert review or psychometric screening
- **Strongly Consider Revision:** Concerns identified in both expert review and psychometric screening

For items that underwent substantial modifications (MOD = 3), original psychometric evidence was considered no longer applicable. These items were therefore classified based on expert judgment only and categorized as:

- **Requires Validation:** No concerns identified in expert review
- **Requires Validation and Revision:** Concerns identified in expert review

This classification framework distinguishes between two complementary dimensions of item evaluation: content validity and contextual appropriateness (as assessed through expert review), and operational item functioning (as assessed through psychometric screening). For items with limited or moderate adaptation, both dimensions are jointly considered to inform suitability for further use. For substantially redeveloped items, the classification reflects the need for empirical validation rather than confirmed measurement limitations.

By separating these dimensions and combining them only at the final classification stage, the approach provides a transparent and reproducible mechanism for identifying items that can be suggested for further use without modification, as well as those requiring revision or validation. A separate listing of country expert-rated policy relevance supports researchers in prioritizing questions for specific LMIC contexts, while additional information on associations with student achievement (only for MOD<3 items) provides further guidance on the empirical relevance of selected constructs.

5.8 SDG Indicator Evaluation

For the current research project, a total of 15 SDG indicators identified as measurable through ILSA context questionnaires in the Analysis of Context Questionnaires in International Large-Scale Assessments (UIS, 2025) were mapped to the indicators constructed in this study (see chapter 5.3.1.2). This mapping was based on the example questionnaire items provided in the UIS analysis document.

All indicators classified as measurable were assigned one score point in the indicator retention rating process (see 5.3.1.3). In a subsequent review conducted by Yanhong Zhang (Chief, Foresight and Innovation) in March–April 2026, the SDG alignment of the retained project indicators was assessed based on the final selection and adaptation of questionnaire items.

In addition to the SDG indicators, a set of supporting context variables required for indicator construction was identified and categorized into two groups. First, disaggregation variables refer to variables necessary for reporting SDG indicators across population subgroups, such as gender, age, or migration background. Second, metric prerequisites refer to variables required for the calculation of specific indicators, for example school enrollment figures needed to compute pupil–teacher ratios (e.g., SDG 4.c.2).

These supporting variables were not included in the scoring process but are documented and discussed in chapter 6.3.3.2.

5.9 AI-Assisted Procedures

AI tools were used to support selected stages of the literature review, comparative analysis of assessment frameworks and questionnaires, development and checking of classification approaches, statistical programming, graphical presentation, and drafting. AI-generated or AI-supported outputs were reviewed against source materials, underlying data, statistical results, and expert judgment before inclusion. AI tools were not used as autonomous decision-makers, and all substantive methodological and interpretive decisions remained with the project team. Further details on the tools used, their purposes, and the verification procedures applied are provided in Appendix A7.

6 Results

6.1 Evidence from the Literature and Areas for Strengthening in IEA ILSA Frameworks

The review presented in Chapter 4 identified a broad set of contextual indicators associated with learning outcomes in LMICs. However, the strength and directness of the supporting evidence vary across indicators. This section summarizes the evidence for indicators that are relevant for understanding learning outcomes in LMIC contexts but are not yet included in context questionnaires used in IEA flagship studies TIMSS and PIRLS. The comparison with regional assessments is used to illustrate that LMIC-focused instruments capture a broader range of contextual factors that remain underrepresented in IEA assessments. LaNA captures a subset of these indicators, while regional assessments such as ERCE, PASEC, SACMEQ, and SEA-PLM cover a broader range, as reflected in Table 6-1.

While remaining indicators are covered by IEA questions, some of the indicators may only have a very limited coverage. For example, the constructs of nutrition and mental health are only represented by an item on how often hunger and tiredness interfere with attention in school. Those indicators therefore should be complemented by collecting more comprehensive information from additional items.

Table 6-1 provides a structured synthesis of these indicators, organized by student, teacher, and school levels. It serves two main analytical purposes. First, it classifies indicators according to the strength of empirical support in the literature, distinguishing between those with strong, direct evidence and those with more context-dependent or indirect associations with achievement. Evidence support is indicated as follows: (+) strong empirical support, referring to indicators with consistent and direct associations with learning outcomes, instructional time, or participation; (±) mixed or context-dependent evidence, referring to indicators whose relationships with achievement are less consistent or more indirect; and (○) limited or no direct empirical support, referring to indicators that are primarily included as contextual or descriptive variables rather than established predictors of achievement. Note that indicators rated as mixed (±) often reflect important contextual mechanisms identified in the literature, but where direct causal relationships with achievement may be less consistently established. Second, the table assesses the extent to which these indicators are represented in LaNA and across the selected regional assessments ERCE, PASEC, SACMEQ, and SEA-PLM.

Indicators such as teacher absenteeism, language of instruction, school infrastructure, learning materials, grade repetition, child work, and health and nutrition (the latter only narrowly covered in the IEA instruments) receive relatively strong support in the literature, as they are directly linked to students' opportunities to learn, instructional time, or participation in schooling. Other indicators, including parental involvement, homework, teacher confidence, contract type, additional paid employment, supervision, school ownership, and household composition, are better understood as mixed or indirect indicators. Their relevance is well established conceptually and in assessment practice, but their relationship with achievement is more context dependent.

This distinction reflects the broader selection logic of the indicator set. Indicators were not chosen solely based on direct empirical associations with achievement, but also for their relevance in LMIC contexts, their feasibility for inclusion in large-scale questionnaires, and their ability to capture meaningful variation in learning conditions.

This is particularly evident for several teacher- and classroom-level factors. While aspects such as teacher preparation, instructional practices, and teacher attendance are well documented, other dimensions—such as employment conditions, contract types, additional paid employment, and teachers’ confidence in instructional practices—are less systematically examined in comparative research. These factors are nevertheless frequently highlighted in regional assessments as relevant factors.

Several student-level indicators, e.g., household composition, housing conditions, or engagement in work outside the household, are not always directly linked to achievement in the same way as traditional socioeconomic background measures. However, they are often used in LMIC contexts as proxy indicators of socioeconomic conditions, particularly where direct measures of income are unavailable.

Certain constructs, such as classroom facilities and teacher job satisfaction, are typically operationalized at the teacher or school level in IEA questionnaires, where partial coverage exists. However, due to the absence of a LaNA Teacher Questionnaire in the 2023 Linking Study, these constructs were categorized at the school level for the purposes of this analysis.

Table 6-1: LMIC-Relevant Contextual Indicators, Evidence support, and Coverage across LaNA and Regional Assessments

LEVEL	# ¹⁸	Indicator	Evidence Support	LaNA	ERCE	PASEC	SACMEQ	SEA-PLM
STUDENT	8	Grade repetition	+		X	X	X	X
STUDENT	10	Household composition (number of people living at home)	±	✓	X	X	X	X
STUDENT	13	Type, material of house	±		(X)	X	X	(X)
STUDENT	14	Parental involvement in learning (incl. help with schoolwork/ Tuition)	±		X	X	X	X
STUDENT	15	Assignment of Homework	±		X	X	X	X
STUDENT	24	House chores/ Work	+		X	X	X	X
STUDENT	25	Commuting to school	+	✓			X	X
STUDENT	26	Language of class instruction	+				X	
TEACHER	5	Professional Teaching Qualification	±		X	X		
TEACHER	6	Subject-Matter Assignments (general /in sel. class)	±		X		X	X
TEACHER	7	Grades/ Classes being taught	±		X	X		

¹⁸ The “#” column lists each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

LEVEL	#18	Indicator	Evidence Support	LaNA	ERCE	PASEC	SACMEQ	SEA-PLM
TEACHER	8	Initial Teacher Education (duration/quality/Experiences)	±		X	X	X	
TEACHER	11	Confidence in teaching (methods)	±		X			X
TEACHER	13	Teacher Absence	+			X	X	
TEACHER	14	Type of contract with school/ Position	±		X	X	X	
TEACHER	15	Additional paid jobs	±		X	X		
TEACHER	16	Instructional Resources and Classroom Infrastructure Inventory	+			X	X	X
TEACHER	22	External Supervision & Inspection	±		X	X	X	
TEACHER	25	Language of instruction	+			X		
SCHOOL	1	Gender	o		X	X	X	X
SCHOOL	2	Age	o		X	X	X	X
SCHOOL	5	Professional development	±		X	X		
SCHOOL	6	School ownership (private/public/religious)	±		X	X	X	X
SCHOOL	7	School type/structure (primary/ secondary/language)	±		X	X	X	
SCHOOL	13	Number of teachers/ Staff	+	✓	X	X	X	X
SCHOOL	19	Classroom facilities	+	✓				
SCHOOL	21	Provision of additional support/ tutoring in reading or math	±			X		X
SCHOOL	25	Shifts	+		X	X	X	
SCHOOL	27	Teacher job satisfaction	±	✓				
SCHOOL	34	Policies on grade retention/ academic support & digital devices	±					
SCHOOL	36	Nutrition, meals provided by school	+	✓			X	

X = indicator covered / (X) = indicator covered only in the parent questionnaire
 (+) strong empirical support/ (±) mixed or context-dependent evidence
 (o) limited or no direct empirical support

6.2 Communalities and Differences across Assessment Frameworks (RQ1)

As described in Section 5.2, the IEA 1989 3-level generic framework was used as an analytical grid to analyze and compare the eight international and regional frameworks selected. Seven frameworks were ultimately included: ERCE, PASEC, PISA-D, PIRLS, SACMEQ, SEA-PLM and TIMSS.

Each of the remaining seven frameworks were analyzed (see Section 5.2 for more details) in terms of the three education levels: system, school/classroom, and student. They were analyzed in terms of any curriculum antecedents: e.g., systems features and conditions; the curriculum contexts, e.g., institutional settings; and the curriculum content, e.g., implemented curriculum.

In this section, the findings from a comparative analysis of the regional frameworks are presented. Thereafter, a comparison of these frameworks with the IEA studies frameworks for PIRLS and TIMSS is presented. This is followed by the identification of shared topic areas, the LMIC specific emphasis, and those dimensions under-represented in IEA frameworks. The section concludes with a proposed new framework.

6.2.1 Comparative Analysis of Regional Frameworks

From a conceptual perspective, regional ILSAs present a distinctive profile. Explicit theoretical or conceptual frameworks are not always articulated in official documentation. SACMEQ, for example, has been described as a pragmatic compromise (Grisay, 1984) between academically oriented outcome measurement and the immediate informational demands of national governments (Grisay & Griffin, 2004). SACMEQ reports tend to emphasize technical assessment frameworks and descriptive analyses rather than explicit theoretical models, although several authors note the influence of IEA's conceptual traditions on SACMEQ design and implementation (Ross & Genevois, 2006). In contrast, PASEC provided an explicit contextual/theoretical framework described in the PASEC 2019 Technical Report, structured across student/family, school/classroom, and system levels and organized around inputs, processes, perceptions, and outputs.

Regarding the conceptualization of regional studies, the presentations of theoretical or conceptual (often used interchangeably) frameworks or models are most explicit in secondary analyses of ILSAs, whereas reports furnished by the bodies conducting the ILSAs appear to embed their theoretical underpinnings into their assessment frameworks, where they are not always explicitly outlined.

However, ILSAs, including regional assessments like SEA-PLM, provide more explicit conceptual or theoretical frameworks/models (as part of the assessment frameworks) in their reporting (UNICEF, 2019).

The assessment frameworks provide various approaches to measuring the environmental and background factors that influence student achievement. While ILSAs like TIMSS and PIRLS focus on global standards, regional and developmental assessments such as PISA-D, SACMEQ, and PASEC offer more targeted measures for LMICs. The following descriptions provide an overview of each of the frameworks included in the analysis:

- PISA for Development (PISA-D): This framework is built upon the Educational Prosperity model, which uses a life-course approach to assess outcomes across six stages of development from conception to age 18. It focuses on “Prosperity Outcomes” (e.g.,

academic performance, health, and well-being) and “Foundations for Success” (e.g., inclusive environments, quality instruction, and material resources). A unique feature is its out-of-school component, which collects data on youth aged 14–16 who are not enrolled in school or are in low grades.

- IEA TIMSS 2023 & PIRLS 2021: These frameworks follow a curricular model based on the IEA’s generic model (see 4.2), addressing the “intended, implemented, and attained curriculum”. They organize contextual data into five broad influences: home, school, classroom, student attributes, and national contexts. They emphasize trend measurement and provide high-quality data to help policymakers improve reading, mathematics, and science education globally.
- SEA-PLM 2019: This regional framework for Southeast Asia identifies four contextual levels: the individual, home/immediate out-of-school environment, school/classroom, and the wider community. It classifies factors as either antecedents (exogenous conditions like SES) or processes (factors directly influencing learning like teaching practice). It includes a specific focus on Global Citizenship.
- SACMEQ: Focused on Grade 6 students in Southern and Eastern Africa, this study aims to assess the conditions of schooling and the backgrounds of students and teachers. It notably includes a module on students’ and teachers’ knowledge about HIV and AIDS. No explicit contextual framework was found to date.
- PASEC: This theoretical framework for Francophone Africa depicts data on student/family, school/classroom/local community, and country/system. It analyzes inputs (e.g., textbooks), transformation processes (e.g., pedagogical practices), and outputs (e.g., student competency levels). It also includes a unique assessment of teachers’ didactic knowledge.
- ERCE / TERCE (LLECE): Used in Latin America, this contextual framework is based on the CIPP model (Context, Input, Process, and Product). The framework includes dimensions of School, apprenticeships, students and families and teachers and teaching. It places a high priority on the school environment, a critical predictor of achievement in the region.

6.2.2 Comparison of Regional Frameworks with IEA Frameworks

Table 6-2 categorizes the system level analyses done on the seven frameworks focusing on their main dimensions and thematic areas. Where there were no explicit contextual/conceptual frameworks found for SACMEQ and PASEC, the dimensions and thematic areas that were included in the data collection for those studies were compiled based on the categories of indicators (and not variables) identified and found in the latest study’s reports as well as historical technical reports. In the other five studies, the main dimensions are included below aligned to the IEA 1989 Framework’s Curriculum Dimensions and the associated Thematic area.

Table 6-2: System-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA studies

Curriculum dimension	Thematic area	TIMSS (2023)	PIRLS 2021	SEA-PLM (2019)	SACMEQ IV (constructed)	PASECs 2019 (constructed)	ERCE (2019)	PISA D
Curricular Content	Intended	Mathematics and Science Curricula	Reading Curriculum	Reading, Writing and mathematics	Reading and mathematics	Reading and Mathematics	Mathematics and reading	Reading, Mathematics and Science
Curricular antecedents	System Features and conditions	Organization of Education System	Organization of Education System	Antecedents and Processes		Structure and areas	Dimension and factors	
Curricular contexts	Institutional Settings	National Contexts	National Contexts	Wider Community Educational system		Characteristics of country, education systems, education zones		

Intended Curriculum: reading, writing, mathematics, and science. The IEA studies are similar in approach to each other but different to the other regional studies. Given the embedded nature of curriculum in IEA studies, the IEA studies are explicit on the Curriculum content Intended level referring to the reading and to the Mathematics and Science Curricula, explicitly. PIRLS refers to the Reading Curriculum in the Primary Grades as well as considering Students with Reading Difficulties explicitly in its framework. This explicitness differs from the other studies where the Curriculum content on the Intended level is implicit, that is, only indirectly reflected in the assessment design and domain selection with studies with available frameworks, or absent where no framework was found.

System features: No explicit features and conditions were mapped for the IEA studies in terms of descriptors in contrast to other studies categorizing system features as antecedents and processes, structure and areas, dimension and factors (these were also not present in SACMEQ and PISA).

Institutional settings: The IEA studies both include National Contexts in their framework explicitly. These describe the constructs sought regarding the Organization of the Education System, namely: System for Pre-primary Education, Age of entry and retention policies, Number of years of school, and Language(s) of instruction. These are common to both studies, but TIMSS also includes Teacher and Principal Preparation. Only two regional studies explicitly seek information on that level: SEA-PLM, with reference to the Wider Community education system, and PASEC, which focuses on Characteristics of country, education systems, and education zones (the national level was not explicit in SACMEQ, ERCE and PISA-D).

Table 6-3 categorizes the school-level analyses done on the seven frameworks focusing on their main dimensions and thematic areas similar to the approach of the system-level analyses above.

Table 6-3: School-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA studies

Curriculum dimension	Thematic area	TIMSS (2023)	PIRLS 2021	SEA-PLM (2019)	SACMEQ IV (constructed)	PASEC 2019	ERCE (2019)	PISA D
Curricular Content	Implemented (Perceived) curriculum	School Emphasis on Academic Success, Mathematics and Science	School Emphasis on Academic Success, Reading instruction	Student Learning process				
Curricular antecedents	Community, School characteristics	School Contexts	School Contexts	Antecedents: School/classroom stated curriculum/resources	School Head Characteristics & Resources	Characteristics of school, class, local community	Schools, Apprenticeships	Educational Prosperity
Curricular contexts	School conditions and processes	School Context	School context	Process: School/ Classroom	School mortality rates of students & teachers (HIV/Aids link)		School Environment	Supporting content: principal views and school information Inclusive environments (school) Supporting content

Implemented (Perceived) curriculum: The IEA studies are similar in approach to each other with both including reference to school's emphasis on academic success and on the subjects being assessed (Mathematics/Science and Reading) and their approach differed from the other regional studies as only SEA-PLM included reference to the curriculum on school-level under Student Learning Process. None of the other studies' frameworks or reference works appeared to refer to this level of curriculum.

Community and School characteristics: Both IEA studies included School Context under which similar antecedents were identified. The studies followed similarly with regards to curricular antecedents focusing on school's characteristics and resources although the terminology and categorization differed from the IEA studies. The regional studies such as SACMEQ, PASEC and SEA-PLM made greater reference to the surrounding community context, accessibility of the school, and access to community resources.

School conditions and processes: The IEA studies both categorized a number of domains from their frameworks under school context relating to conditions and processes at school. This was one domain where the IEA studies differed from each other in their frameworks with PIRLS including time spent on language and reading instruction under "School Context" and a second topic on Covid-19 pandemic whereas no similar reference was found for Mathematics and Science under this domain. The regional studies varied greatly both from the IEA studies as well as each other. SEA-PLM focused significantly on school conditions and processes while

SACMEQ and ERCE had less emphasis. PISA's related topics on school conditions and processes were regarded as "Supporting content" although there was an element of the framework focusing on inclusive environments, terminology not found in the other studies. Significantly, SEA-PLM, SACMEQ and PASEC all included topics on School health programs, extending in some cases to mortality rates and information related to HIV/Aids (SACMEQ and PASEC). Other examples include teachers' negative behavioral traits such as truancy and substance abuse (found in SACMEQ and PASEC), which were not included in the IEA studies.

Table 6-4 categorizes the class-level analyses done on the seven frameworks, focusing on their main dimensions and thematic areas as was done with the system and school-level.

Table 6-4: Class-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA Studies

Curriculum dimension	Thematic area	TIMSS (2023)	PIRLS 2021	SEA-PLM (2019)	SACMEQ IV (constructed)	PASEC 2019	ERCE (2019)	PISA D
Curricular Content	Implemented (operational curriculum)	Emphasis on science enquiry Classroom Contexts Mathematics and Science Instruction Classroom Assessment	School Emphasis on Reading Instruction Classroom Reading Instruction Teachers Develop Students' Reading Comprehension Skills and Strategies Classroom Assessment			teachers' teaching practices and perceptions	Teachers and teaching	Quality instruction, Learning Time teacher's teaching multi-grade classrooms, holds multiple teaching jobs or works other jobs in addition to teaching
Curricular antecedents	Teacher characteristics	Teacher Characteristics	Teachers' Preparation	Antecedents: general professional background of teachers	Teachers' characteristics	Teachers' personal and professional characteristics	Teachers and teaching	Supporting content (teachers background and perceptions), teachers' age and sex
Curricular contexts	Classroom conditions and processes	Classroom Contexts/ Information Technology in the classroom Classroom Climate	Classroom Contexts (e.g. classroom libraries)/ Information Technology in the classroom Classroom Climate Classroom Management	Antecedents:(School) / Classroom provision of text books, teaching aids classroom management, practices and activities policies and procedures for literacy development and numeracy development (remedial classes)	Essential & Desirable School resources (classroom)	Classroom infrastructure, equipment, functioning and resources Teacher absenteeism, student truancy and classroom learning opportunities Classroom management, Professional Responsibilities	Teachers and teaching	material resources inclusive environment (classroom)

Implemented (operational curriculum). Both IEA studies included School emphasis in their respective subject focus; however, while this was an explicit dimension in the PIRLS framework, this was not the case in the TIMSS Framework where the emphasis was less prominent. PIRLS appeared to place greater focus on this level of the curriculum. The regional studies differed considerably on this dimension with broad reference to “Teachers and teaching” (PASEC and ERCE), Quality instruction (PISA-D) or no reference (SEA-PLM and SACMEQ). However, PISA-D also included a specific factor “Learning Time”.

Teacher characteristics: in terms of Classroom curricular antecedents, the IEA studies also differed from each other in approach, terminology “Teacher Characteristics” (TIMSS) and Teacher Preparation (PIRLS) and contents. However, all the regional studies included aspects of teacher characteristics in terms of demographics, qualifications, and experience, albeit structured slightly differently. In the case of PISA-D, this was seen as “supporting content”.

Classroom conditions and processes: PIRLS extended its framework relative to TIMSS with regards to classroom conditions and processes including the topic of classroom management explicitly. The regional studies varied greatly from the IEA studies with ERCE generally referring to “teaching” and PISA-D “material resources” and “inclusive environment” to SEA-PLM elaboration on teaching aids, classroom management practices and activities from the students’ perceptions as well as policies and procedures for literacy and numeracy development. Similarly, in the reconstructed framework for SACMEQ and the framework for PASEC, classroom conditions received strong emphasis. However, notably in SACMEQ and PASEC, there was a focus on “essential” and “desirable” resources (SACMEQ terminology) and an implicit assumption that not all students would have their own desk, chair or textbooks. This was quite different from the IEA studies. Another difference was the probing of teachers’ absenteeism and student’s truancy also not present in the IEA studies. The IEA studies also included Information technology and its use in the classroom, which was not explicit in the other frameworks.

Table 6-5 categorizes the student-level analyses done on the seven frameworks focusing on their main dimensions and thematic areas as was done with the system and school level.

Table 6-5: Student-Level Curriculum Dimensions and Thematic Areas Addressed in Regional Frameworks Compared with IEA Studies

Curriculum dimension	Thematic area	TIMSS (2023)	PIRLS 2021	SEA-PLM (2019)	SACMEQ IV (constructed)	PASEC 2019	ERCE (2019)	PISA D
Curricular Content	Attained (experiential and learned)	Achievement in Mathematics and Science	Achievement in Reading Literacy	Outcome: Mathematical, reading and writing literacy, and global citizenship	Achievement in Reading and Mathematics	Achievement in Reading and Mathematics	Achievement in Mathematics and Reading	Prosperity Outcomes: Academic Performance, Prosperity outcomes: Health and well-being
Curricular antecedents	Student Background characteristics	Home Contexts, Student Attributes	Home Contexts	Home environment / Family Background Antecedent: Home environment: Antecedents: Student Characteristics	Pupil characteristics	Characteristics of students and family environment	Students and their families	Equality, Equity & Access (Home)
Curricular contexts	Student (classroom) behaviors	Attitudes Toward Mathematics and Science/ Valuing Mathematics and Science	Student Reading Attitudes	Process: Home environment Process: Student Learning Process: Students' enjoyment of reading, writing and mathematics	Health Knowledge (HIV and AIDS)	Appetite for learning,	involvement learning dispositions exposure time to teaching	Equality, Equity & Access (student) Prosperity outcomes: Attitudes towards school & learning

Attained (experiential and learned): The IEA studies focused on the subject specific achievement in terms of the attained curriculum. While SAQCMEQ, PASEC and ERCE had similar curricular content, SEA-PLM and PISA-D differed. SEA-PLM included writing literacy as well as global citizenship as outcomes in addition to mathematics and reading literacy. PISA-D included prosperity outcomes in the form of both academic performance but also health and well-being; the latter not included by any other study including the IEA studies.

Student Background characteristics: Curricular antecedents on student level differed between the IEA studies with a greater emphasis by PIRLS on “learning” for example Home Resources for Learning/Environment for Learning compared with TIMSS. Parental expectations for Further Education were included in TIMSS and not in PIRLS, and there was a focus on Information Technology and Digital devices in TIMSS. The PIRLS framework focused on Home contexts while TIMSS included both Home Contexts and Student Attributes on domain level illustrating perhaps the difference in target age groups. The regional studies varied with three studies referring to “families” or “family environment” or “Family background” and two studies referring to “home” although SEA-PLM had also included “family background” and PISA-D positioned the “home” within an Equality, Equity and Access dimension and SACMEQ referred to “pupil characteristics”. While there were overlaps on topic level, the different emphasis and terminology used is indicative of the different conceptual approaches taken.

Student (classroom) behaviors: The IEA studies both had dimensions of attitudes toward the subject they were assessing, but TIMSS also included students “valuing” the subject as a dimension on the framework. The regional studies all included some dimension of attitude toward learning or school generally although terminology varied. SACMEQ appeared unique in its inclusion around knowledge and conceptions regarding HIV/Aids and ERCE in including the exposure time to learning in its student-level dimension of its framework.

6.2.3 Similarities and Differences

The analysis done on the frameworks at domain and thematic areas attempted to identify similarities and differences between the IEA studies and the regional studies. Notwithstanding that each study had a specific purpose and that each of the studies would have been conceptualized accordingly and with different target populations. Nonetheless, some interesting similarities and differences were found (see Table 6-6).

Each of the framework’s domains/topics were analyzed to the greatest extent possible (see section 5.2 for details). In Table 6-6, the domains and topics in studies with explicit contextual frameworks were coded as being “Present in framework at domain/topic level”. Where the domain was not explicit, but its associated constructs were described within the framework document, they were coded as “implicit in the framework”. Where the study did not have a contextual framework, but the topics could be construed from the categorizations and classifications presented in the international or technical reports, they were coded as “implicit and derived” from reports. Where the domains or topics were not explicit or construed, they were coded as “Not found on domain or topic level or absent”. This should not be interpreted as the complete absence of any data for that dimension, but rather that the visibility and emphasis allotted by the studies to specific domains and topics on the Framework level was limited. The analyses were based upon the original IEA frameworks (see sections 4.2 and 5.2 for more details).

Table 6-6 provides a consolidated summary of the detailed framework comparisons presented in

Table 6-2 through Table 6-5. Whereas the preceding tables examined the conceptual structure and contextual domains of each assessment framework individually and comparatively across curriculum dimensions, Table 6-6 serves as a synthesis overview highlighting the presence, emphasis, and relative alignment of major domains across the reviewed international and regional large-scale assessments.

The table is organized according to the three overarching curriculum dimensions derived from the IEA framework tradition: curricular content, curricular antecedents, and curricular contexts. Within these dimensions, the frameworks are compared across thematic areas and levels of the education system, including system, school/classroom, home, and student levels. This structure enables the visualization of how different assessments conceptualize learning processes, contextual influences, and educational outcomes, as well as where frameworks converge or diverge conceptually.

Table 6-6: Overview of the Domains and Topics Found Across Curricular Domains on System, School, Class and Student Levels in Selected International and Regional Studies

Level	Curricular Content	Curricular antecedents	Curricular contexts	TIMSS (2023)	PIRLS 2021	SEA, PLM Southeast Asia Primary Learning Metrics (2019)	SACMEQ IV	PASEC 2019	ERCE Regional Comparative and Explanatory Study (2019)	PISA- D 2018
System	Intended	System Features and conditions	Institutional Settings	Mathematics and Science Curricula	Reading Curriculum	Reading, Writing, Mathematics, Global Citizenship	Reading, Mathematics, HIV/Aids Knowledge	Reading and mathematics	Mathematics, Science and Reading	Reading, Mathematics and Science
				Organisation of Education System	Organisation of Education System	Antecedents and Processes		Structure and areas	Dimension and factors	
				Covid_19 Pandemic				Characteristics of country, education systems, education zones		
School or Classroom	Implemented	Community, School and teacher characteristics	School and classroom conditions and processes	School	School	School/Classroom		Characteristics of school, class, local community	Schools, Apprenticeships	
						Student Learning process				
				School Contexts	School Contexts	Antecedents: School/classroom stated curriculum/ resources	School Head Characteristics & Resources	Characteristics of school, class, local community	Schools, Apprenticeships	Educational Property
						Process: School/Classroom			School environment	
(School or) Classroom	Attained	Student Background characteristics	Student classroom behaviours	Emphasis on science enquiry	School Emphasis on Reading Instruction				Teachers and teaching	
									Teachers and teaching	
				Instruction Limited by Student Attributes		Antecedents:(School) / Classroom				
	Attained	Student Background characteristics	Student classroom behaviours	Achievement in Mathematics and Science	Achievement in Reading Literacy	Outcome: Mathematical, reading and writing literacy, and global citizenship	Achievement in Reading, Mathematics, HIV/Aids	Achievement in Reading and Mathematics	Achievement in Mathematics, Science, Reading and Writing	Prosperity Outcomes: Academic Performance, Prosperity outcomes: Health and well-being
				Home Contexts, Student Attributes	Home Contexts	Home environment / Family Background	Pupil characteristics	Characteristics of students and family environment	Students and their families	Equality, Equity & Access (Home)
						Process: Home environment				Equality, Equity & Access (student)

	Dark green – present in framework at domain/ topic level
	Light green – implicit in framework (construct/ variable level)
	Yellow – implicit as no framework, derived from reports
	White – not found at domain level or topic level or absent

6.2.3.1 Communalities Across Frameworks

All frameworks utilize a multi-level structure (individual, home, school, and system levels) to explain variations in learning. Nearly every framework considers Socioeconomic Status (SES) a primary determinant of student performance and uses home possessions as a proxy for wealth. Most frameworks are moving toward measuring student well-being and social-emotional skills alongside cognitive outcomes.

6.2.3.2 Differences Across Frameworks

PISA-D is distinguished by its life-course/developmental perspective and its specific focus on out-of-school youth, whereas TIMSS and PIRLS remain strictly school- and grade-based. Regional assessments like SACMEQ and PASEC place more emphasis on basic material inputs (e.g., toilets, electricity, and clean water) that are often universal in high-income countries but show high variance in LMICs. SEA-PLM is unique in its integration of a Global Citizenship literacy domain.

6.2.3.3 Further Observations

In addition to the communalities and differences, some additional observations were made. The pervasiveness of the original IEA framework is evident as underpinning both IEA studies with the domains being either explicitly or implicitly visible across the full length of the framework. Secondly, the absence of an explicit framework for SACMEQ may explain the absence of the domains/topics able to be aligned to the IEA framework. The focus on student level and school level across all the studies is visible. While teachers are critical actors in an education system, they are not prominent in the frameworks and serve more as “supporting content” conceptually across the frameworks.

6.2.3.4 LMIC-Specific Emphases

Common to the LMIC frameworks were assumptions about the impact of poverty on schooling and classroom conditions where more information was sought about the basic conditions and resources in the classroom (e.g.: chairs, desks for each child, teaching aids for teachers) as well as what is deemed “essential” vs “desirable” resources. Emphasis on the nature of community contexts and location of the school relative to the student’s home was also prevalent among the LMICs.

6.2.3.5 Dimensions Under-Represented in IEA Frameworks

The IEA studies differed from the regional studies in their assumptions about what was expected to be in an average classroom and school and therefore was not included conceptually in the original IEA framework. Similarly, the regional studies in their approach interrogated the schools and teachers about the behavior and presence of the teachers in a manner not done in the IEA studies. The IEA studies did not include outcomes regarding well-being and nor that of the teachers.

To better reflect the educational realities of LMICs, the review identified several areas emphasized in regional assessments and PISA-D that could strengthen the contextual relevance and applicability of IEA frameworks:

- **Basic/Essential School Services:** As mentioned earlier the IEA frameworks tend to omit basic infrastructure (e.g.: variables like electricity, running water, and sex-disaggregated toilets) which are essential prerequisites for learning in many LMIC schools.

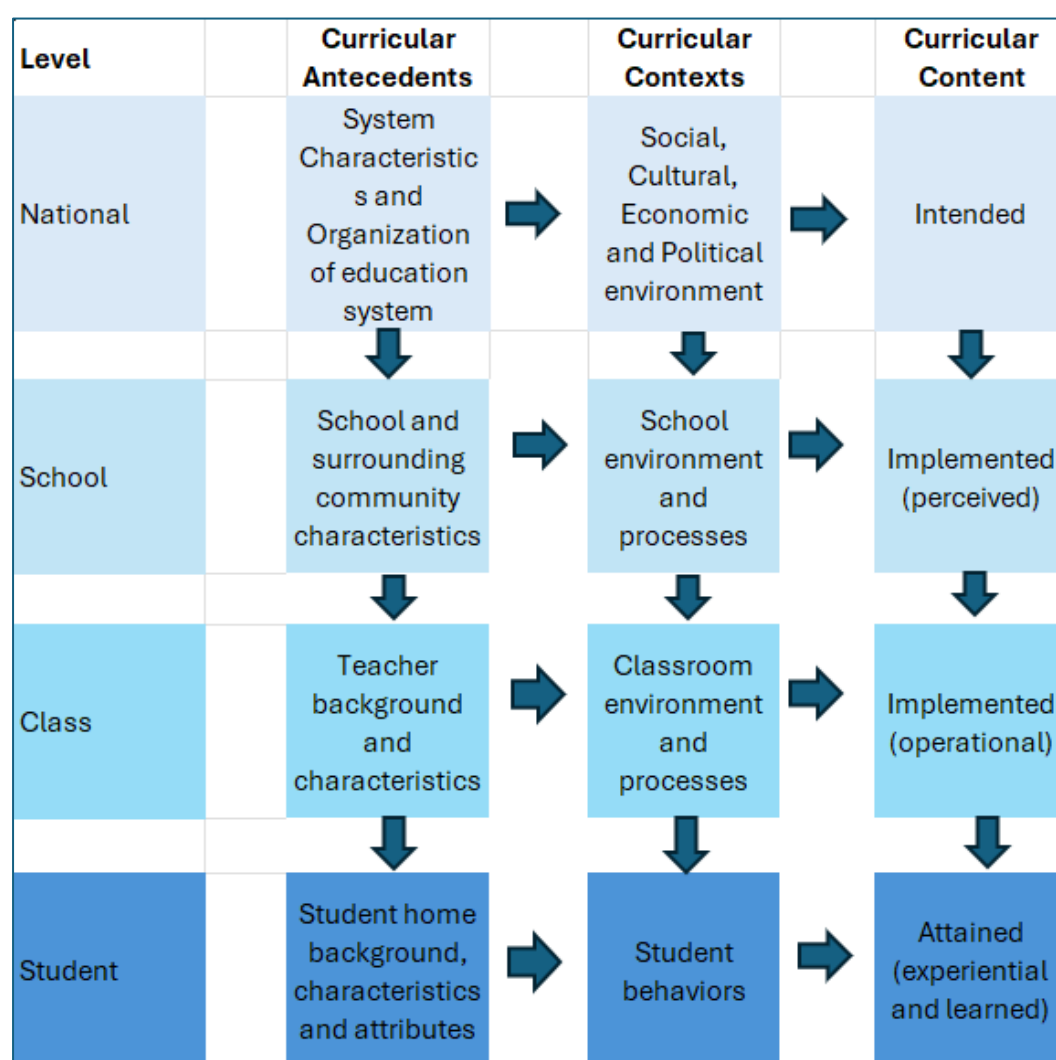
- Detailed socioeconomic background and Poverty Measures: More sensitive poverty indicators, such as dwelling construction materials (roof/wall/floor condition) and food security/nutrition (e.g., school meal availability), are needed.
- Teacher Labor Realities: frameworks should address teacher absenteeism, payment regularity, and secondary income activities (e.g., teachers working multiple jobs or as private tutors), which significantly impact instruction in LMICs.
- Language of Instruction Complexity: While IEA asks about home language, the nuance regarding the transition point between mother-tongue instruction in early grades and the switch to a national/official language in later primary years, which is a major barrier for many LMIC students.
- Health, wellbeing and Nutrition: Explicit questions on chronic illnesses, available nutrition/meals, or hunger at school are directly related to low performance in LMICs than the attitudinal factors often prioritized in HIC-focused frameworks.

6.2.3.6 Toward a New Model

Given the origin of the generic IEA Framework and the significant time that has passed since its design, and the changing nature of the participating countries in the IEA studies, the outcomes of the framework analysis suggest some changes are warranted as listed above. While upholding the continued focus on the curriculum, the changes on school and classroom level are required to account for the new realities regarding the perceived and operational curricula, teachers, teaching and learning processes and classroom environments and conditions. The separation on student-level of the home environmental factors and student attributes and behaviors suggests a clearer separation and focus on framework level.

Therefore, the model shown in Table 6-7 is suggested to provide the basis for defining explicit factors on four separate levels.

Table 6-7: Proposed Framework Based on Outcomes of Analyses



The new framework extends some of the previous concepts proposed in the original IEA Generic framework. It consolidates long lists of variables into process-oriented constructs. The new framework preserves the distinction between inputs (resources, characteristics) and processes (leadership, teaching) are preserved but at a higher level of abstraction. The proposed framework distinguishes between four explicit levels of the education system instead of the previous three levels. The separation of school and classroom level characteristics and processes and levels of curricula provide an in-depth intentional focus on levels of variation within systems and schools and allows conceptually for the distinction between the perceived and operational curriculum at school and classroom levels within schools and the operationalization of these.

On system level, the new framework shifts from cataloguing system features to explaining system-level variation, aligning with explanatory system models and statistical developments since 1980s. The curriculum is no longer foregrounded as a standalone object; instead, it conceptually embeds school-based curriculum, instructional leaderships and school educational emphases within the perceived curriculum. Therefore, the analytical emphasis moves from what exists to how system arrangements function and differ.

On school and classroom levels, the principal and teacher characteristics are embedded within leadership, teaching, and classroom antecedents. The shift moves from inventory-style reporting to analytic constructs capable of explaining between-school and between-classroom variation. The school and classroom are treated as discrete, but interacting systems, rather than as parallel lists of features. The school and classroom level topics are collapsed and reframed now in bespoke levels.

On student level, the new framework adopts a learner-centered, holistic model. The student characteristics, home background, attitudes, behaviors, and engagement are integrated as attributes and dispositions, rather than isolated variables. The attained curriculum becomes more clearly aligned with achievement and experiential learning. The student behaviors are no longer marginal descriptors as in the original generic model, but are central explanatory constructs tied to engagement and learning outcomes.

By reframing dimensions of the 1989 IEA generic model into broader, integrated constructs, it signals a conceptual consolidation and modernization of the framework. The new framework strives to reflect the developments since 1989 building on the previous research traditions of the IEA to be strongly explanatory and theory-oriented. It is designed to be integrative and construct-driven. The focus of the new framework is on systems, processes and interactions across levels providing a holistic generic model to guide future developments that are contextually sensitive.

6.3 Indicator Identification and Alignment Across Assessments (RQ2)

This section addresses Research Question 2, which examines how contextual indicators at the end of primary education align across international and regional large-scale assessments. Building on the framework comparison and indicator development process outlined in Chapter 0, the following analysis presents the identified indicator set and examines the degree of conceptual and operational convergence across programs.

The section first summarizes the mapping of identified indicators to the analytical framework and the resulting retention patterns. It then presents the retained indicators and their organization across student, teacher, and school levels. It then examines the retained indicator set and the extent of conceptual alignment across assessments, highlighting areas of strong convergence as well as domains where coverage is more limited or differentiated.

6.3.1 Summary of Indicator Mapping to the Analytical Framework

Table 6-8 presents the results of the mapping between the analytical framework developed for this study and the contextual domains and indicators identified across the reviewed international and regional large-scale assessments. The table is structured according to the overarching curriculum dimensions of curricular content, curricular antecedents, and curricular contexts, and further organized by thematic areas and analytical levels (system, school/classroom, home, and student).

Table 6-8: Indicator-Matching with Project Framework

Level	Category	Retained	Not Retained	Total
School	Antecedents (School characteristics & community)	21	13	34
School	Contexts (School environment & processes)	14	25	39
School	Content (Implemented – perceived)	1	2	3
Class	Antecedents (Teacher background & characteristics)	15	15	30
Class	Contexts (Classroom environment & processes)	10	5	15
Class	Content (Implemented – operational)	2	10	12
Student	Antecedents (Student home background & attributes)	18	11	29
Student	Contexts (Student classroom/learning behaviors)	3	7	10
Student	Content (Attained – experiential & learned)	6	6	12
	TOTAL (184 indicators, National excluded)	90	94	184

For each framework domain, Table 6-8 indicates the number of corresponding indicators that could be identified in the reviewed assessment instruments and whether these indicators were retained for the subsequent analysis and questionnaire development process. “Kept” indicates that the domain was regarded as conceptually relevant for the present study and that corresponding questionnaire items from international or regional assessments were identified and selected for further analysis. Domains or indicators not marked as retained were either insufficiently represented across studies, considered outside the scope of the present framework, or not operationalized consistently enough to support meaningful comparative analysis.

Table 6-8 therefore serves as an intermediary synthesis step between the conceptual framework comparison and the operationalization phase of the study. It illustrates how the broad theoretical and conceptual domains identified across frameworks were translated into analyzable indicator areas and ultimately into concrete questionnaire content.

Overall, the results demonstrate a strong alignment between the analytical framework and the contextual structures of the reviewed ILSAs, particularly for domains related to student background, home learning environments, school conditions, instructional processes, and student attitudes. The distribution of indicators across the framework dimensions further illustrates that some thematic areas were represented by a substantially larger number of indicators than others. Domains related to student and home background characteristics, instructional processes, and school contexts were strongly represented, whereas fewer indicators were identified for broader developmental outcomes and certain context-specific dimensions.

The detailed assignment of each indicator to its corresponding framework domain is presented in Appendix A4.

6.3.2 Overview of Retained Indicators

Following the framework-level mapping presented above, the indicators retained for further analysis through the structured selection process described in Section 5.3 are presented and discussed at the student, teacher, and school levels.

Indicator retention was guided by cross-assessment coverage, alignment with SDG-related monitoring needs, and partner priorities. Indicators meeting predefined thresholds were retained, with professional judgment applied to preserve strategically important constructs in LMIC contexts. The resulting indicator set reflects constructs that are either shared across multiple assessment programs or considered substantively policy relevant.

6.3.2.1 Student-Level Indicators Relevant in LMIC Contexts

At the student level, a total of 27 indicators were retained following the indicator creation and selection procedure (see

Table 6-9). The distribution of coverage scores reveals a stable core of widely shared constructs, alongside a set of more context-specific, regionally relevant indicators. Overall, shared indicator coverage is strongest for demographic characteristics and home environment factors, while greater variation emerges in constructs related to school learning conditions, psychosocial well-being, and structural constraints affecting learning in LMIC contexts.

Table 6-9: Student Level Indicators

# ¹⁹	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
1	Gender	✓	X	X	X	X	X	X	X	X		8
2	Age	✓	X	X	X	X	X	X	X	X		8
3	Language experience	✓	X	X	X	X	X	X	X	X		8
4	Immigration Status		X	X					X	X	X	5
5	Parents' literacy		(X)		X			X	X	X		5
6	Parents' Highest formal education		(X)	X		X	X	X	X	X		7
7	Participation in formal pre-primary education		(X)	X	X	X	X	X	X	X		8
8	Grade repetition			X	X	X	X		X	X		6
9	Absenteeism	✓	X	X		X			X			4
10	Household composition (number of people living at home)	✓		X	X	X	X	X	X		X	7

¹⁹ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

#19	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
11	Number of books in the home	✓	X	X	X	X	X	X	X	X		8
12	Household possessions and services	✓	X	(X)	X	X	X	X	X	X	X	9
13	Type, material of house			(X)	X	X	(X)		X	X		6
14	Parental involvement in learning (incl. help with schoolwork/ Tuition)			X	X	X	X	X	X	X		7
15	Assignment of Homework			X	X	X	X					4
16	Enjoyment of reading/ mathematics	✓	X		X	X						3
17	Confidence in reading/ mathematics	✓	X	X								2
18	Nutrition, meals	✓	X		X	X	X	X				5
19	Physical and mental health (incl. Anxiety, Depression)	✓	X		X			X	X		X	5
20	Well being at school	✓	X	X	X	X	X		X			6
21	Classroom climate		X	X			X		X			4
22	Experience of bullying/ harassment		X	X					X	X	X	5
23	Access to books from community/school/class library		X			X						2
24	House chores/ Work			X	X	X	X		X			5
25	Commuting to school	✓				X	X		X			3
26	Language of class instruction					X			X			2
27	Perception of Instruction	✓	X	X	X		X		X		X	6

X = indicator available in student questionnaire

(X) = indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

✓ = unlike the other assessments, LaNA questionnaires were not included in the scoring procedure for indicator selection. Therefore, indicator coverage in LaNA is denoted by a checkmark (✓) rather than an X

6.3.2.1.1 Widely Shared Indicators Across Assessments

A set of nine indicators demonstrates high coverage across the different assessment initiatives reviewed (scores of 7–9), including alignment with SDG monitoring priorities and partner requests. These include gender, age, language experience, parental education, participation in pre-primary education, household composition, number of books at home, household possessions and services, and parental involvement in learning. These constructs are consistently represented across both IEA and regional assessments and form the stable core of student background measurement.

This shared coverage reflects broad international consensus on the importance of sociodemographic characteristics and socioeconomic status proxies for explaining variation in learning outcomes. All of these indicators align closely with SDG monitoring priorities, particularly regarding equitable access to education and early childhood participation.

6.3.2.1.2 Moderate Shared Indicator Coverage and Context-Sensitive Constructs

A second group of 14 indicators (scores of 4–6) shows moderate but consistent representation across multiple regional assessments. These include immigration status, parents' literacy, grade repetition, absenteeism, type and material of housing, assignment of homework, nutrition and meals, physical and mental health, well-being at school, classroom climate, experience of bullying, commuting to school, house chores, and perception of instruction.

Several of these constructs are less prominently covered within the IEA core frameworks but appear repeatedly in regional programs. This is particularly evident for indicators capturing structural and environmental conditions that are especially relevant in LMIC contexts, such as material living standards, homework practices, child labor responsibilities, and health conditions.

6.3.2.1.3 Expert-Retained Low-Coverage Indicators

A smaller number of indicators exhibit limited cross-study coverage (scores ≤ 3), including confidence in reading or mathematics, access to library resources, and language of instruction. These indicators did not reach the predefined retention threshold based on coverage alone but were retained through expert override due to their substantive relevance, particularly for LMIC contexts.

Confidence in reading and mathematics has been implemented in LaNA and has demonstrated robust measurement properties in LMIC applications, supporting its continued inclusion. Similarly, language of instruction and access to books are highly relevant in multilingual and resource-constrained contexts, where language mismatch and limited access to reading materials represent significant barriers to learning. Although these indicators are less consistently represented across assessments, they can be considered valuable optional components in specific regional settings.

6.3.2.1.4 Indicators Beyond the TIMSS/PIRLS Framework and Positioning of LaNA

Six indicators in total (and nine indicators when referring to the student questionnaire only) with substantial representation in regional assessments are not covered in the TIMSS/PIRLS student questionnaires. These include grade repetition, household composition, type and material of housing, assignment of homework, house chores, and language of instruction. Their repeated inclusion includes dimensions of student experience that are particularly relevant in LMIC contexts but remain underrepresented in the IEA core framework.

Two further indicators — nutrition and meals, and physical and mental health — are covered in LaNA, TIMSS, and PIRLS only through a single item asking how often hunger or tiredness interferes with attention in school, hence this question captures only a narrow facet of each construct.

Among these, household composition is already covered within LaNA, while the remaining indicators fall largely outside its current scope. This pattern suggests that while the TIMSS/PIRLS framework captures widely agreed-upon demographic and socioeconomic constructs, regional assessments contribute additional context-sensitive indicators that reflect common LMIC realities.

6.3.2.2 Teacher-Level Indicators Relevant in LMIC Contexts

At the teacher level, a total of 27 indicators were retained following the indicator creation and selection procedure (see Table 6-10). Given the absence of an AMPL questionnaire and the correspondingly lower maximum score (8), indicator groupings were adjusted accordingly by setting the retention threshold to a score of 3.

Overall, coverage is strongest for basic demographic and qualification-related indicators, while greater variation emerges for constructs related to instructional practices, working conditions, and school environment.

Table 6-10: Teacher Level Indicators

# ²⁰	Indicator	In LaNA	TIPi	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
1	Teacher Gender		X	X	X	X	X		X	X		7
2	Teacher Age		X	X	X	X	X		X			6
3	Years of Teaching		X	X	X	X	X		X			6
4	Highest level of education		X	X	X	X	X		X	X	X	8
5	Professional Teaching Qualification			X	X					X		3
6	Subject-Matter Assignments (general /in sel. class)			X		X	X					3
7	Grades/ Classes being taught			X	X				X			3
8	Initial Teacher Education (duration/quality/Experiences)			X	X	X			X	X		5

²⁰ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

#20	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
9	Area(s) of Post-secondary Study / Teacher		X		X		X			X		4
10	Professional development (duration, quality)		X	X		X	X		X	X		6
11	Confidence in teaching (methods)			X			X					2
12	Job satisfaction (incl. General working conditions)		X	X	X	X	X		X			6
13	Teacher Absence				X	X			X			3
14	Type of contract with school/ Position			X	X	X			X		X	5
15	Additional paid jobs			X	X				X			3
16	Instructional Resources and Classroom Infrastructure Inventory				X	X	X		X			4
17	School- or Classroom library and use		X	X		X				X		4
18	EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)		X	X	X							3
19	EAS_Parental Support		X	X			X		X			4
20	SOS_Personal Safety (including bullying)		X	X	X							3
21	SOS_Student Behavior & Respect		X	X			X		X			4
22	External Supervision & Inspection			X	X	X			X			4
23	Class size/ Multi-single grade classroom		X		X				X		X	4
24	Teaching conditions/ Teaching Constraints		X	X			X					3

#20	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
25	Language of instruction				X				X		X	3
26	Instructional time (Language/ Reading or mathematics)		X		X	X						3
27	Instructional activities (general, Language/ Reading, Math)		X	X			X		X			4

X = indicator available in teacher questionnaire

6.3.2.2.1 Widely Represented Indicators Across Assessments

A set of six indicators demonstrates high coverage across the different assessment initiatives reviewed (scores of 6–8), including alignment with SDG monitoring priorities and partner requests. These include teacher gender, teacher age, years of teaching experience, highest level of education, professional development, and job satisfaction. These constructs are consistently represented across both IEA and regional assessments and form the core of teacher background and qualification characteristics.

This pattern reflects a broad international consensus on the importance of teacher qualifications and experience as key determinants of instructional quality and student learning outcomes. The importance of formal education levels and access to professional development is also reflected in their alignment with SDG indicators related to teacher qualification and teacher training.

6.3.2.2.2 Moderate Coverage and Context-Sensitive Professional Constructs

A larger second group of altogether 20 indicators (scores of 3–5) shows moderate but consistent representation across multiple assessments. These include professional teaching qualification, subject-matter assignments, grades or classes being taught, initial teacher education, area of post-secondary study, teacher absence, type of contract, additional paid jobs, instructional resources and classroom infrastructure, school- or classroom library and use, instructional and curricular leadership, parental support, personal safety, student behavior and respect, external supervision and inspection, class size, teaching conditions, language of instruction, instructional time, and instructional activities.

These constructs extend beyond basic teacher characteristics and capture dimensions of teaching practice, working conditions, and school-level environments. Ten out of 20 of these indicators are included in IEA frameworks, while the remainder are more specific to regional assessments. Overall, coverage is heterogeneous and tends to be broader in scope in regional programs, particularly PASEC and ERCE.

6.3.2.2.3 Expert-Retained Low-Coverage Indicators

Only one indicator (confidence in teaching) was retained despite very limited cross-study coverage (score = 2), and thus below the predefined retention threshold for the teacher level.

Its inclusion was justified based on its conceptual relevance for instructional quality. Confidence in teaching is closely linked to teacher self-efficacy and remains an important construct for understanding teaching practices and student learning outcomes.

6.3.2.2.4 Indicators Beyond the TIMSS/PIRLS Framework

Eleven indicators with notable representation in regional assessments are not fully covered within the TIMSS/PIRLS teacher questionnaires: Professional teaching qualification, subject-matter assignments, grades or classes being taught, initial teacher education, confidence in teaching, teacher absence, type of contract, additional paid jobs, instructional resources and classroom infrastructure, external supervision and inspection, and language of instruction.

Many of these indicators point to more formal and organizational aspects of the teaching profession, which are particularly salient in LMIC contexts, where staffing patterns, contractual arrangements, resource constraints, and administrative structures directly shape classroom practice.

6.3.2.3 School-Level Indicators Relevant in LMIC Contexts

At the school level, a total of 36 indicators were retained following the indicator creation and selection procedure (see Table 6-11). The coverage pattern shows a stable core of widely represented institutional and structural characteristics, alongside a broader set of context-sensitive indicators capturing school resources, organization, and learning environments. Coverage is strongest for basic school characteristics and resource availability, while greater variation emerges in constructs related to school climate, governance, and operational conditions, particularly in LMIC contexts.

Table 6-11: School Level Indicators

# ²¹	Indicator	In LaNA	TPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
1	Gender			X	X	X	X	X			X	6
2	Age			X	X	X	X	X			X	6
3	Years of service, experience in education		X	X	X	X	X	X				6
4	Highest level of education		X	X	X	X	X	X			X	7
5	Professional development			X	X						X	3
6	School ownership (private/public/religious)			X	X	X	X	X	X		X	7

²¹ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

#21	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
7	School type/structure (primary/secondary/language)			X	X	X			X			4
8	School location (urban/rural)	✓	X	X	X	X	X	X	X	X		8
9	SES Backgrounds of students		X	X		X			X	X		5
10	School size (total enrollment of boys and girls and at target grade)	✓	X	X		X	X	X	X	X		7
11	Enrollment & number of classes in target grade		X			X	X	X	X	X		6
12	Language of test/ language of instruction		X				X	X		X		4
13	Number of teachers/ Staff	✓		X	X	X	X	X	X	X		7
14	School Educational Resources except books (and affection by shortage)	✓	X	X							X	3
15	School human resources (and affected by shortage)	✓	X	X			X					3
16	School Library (equipment and use)	✓	X	X		X	X	X				5
17	Internet		X	X		X	X	X		X	X	7
18	School facilities (incl. Toilets SDG)	✓	X	X	X	X	X	X	X	X		8
19	Classroom facilities	✓										0
20	Availability of textbooks	✓	X	X			X	X	X			5
21	Provision of additional support/ tutoring in reading or math				X		X	X	X			4

#21	Indicator	In LaNA	TPII	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
22	Digital learning resources	✓	X	X		X	X			X		5
23	School Year Structure (beginning/end of school year, operational days)	✓	X		X	X					X	4
24	School Year Structure (instructional time per day)	✓	X	X			X					3
25	Shifts			X	X	X						3
26	Days school closed (Covid, strikes, etc.)	✓	X		X	X			X		X	5
27	Teacher job satisfaction	✓										0
28	EAS_Instructional and Curricular Leadership and Staff Cohesion	✓	X	X			X					3
29	EAS_Parental Support	✓	X	X			X				X	4
30	EAS_Student Engagement and Peer Culture	✓	X				X					2
31	SOS_Personal Safety / BULLYING (incl. Teachers)	✓	X	X		X	X		X	X		6
32	SOS_School rules and enforcement	✓	X			X						2
33	SOS_Student Behavior & Respect	✓	X	X		X	X				X	5
34	Policies on grade retention/ academic support & digital devices								X			1
35	Teacher Absenteeism, Teacher availability, Issues at school	✓	X		X	X			X			4
36	Nutrition, meals provided by school	✓				X		X	X			3

X = indicator available in school questionnaire

✓ = unlike the other assessments, LaNA questionnaires were not included in the scoring procedure for indicator selection. Therefore, indicator coverage in LaNA is denoted by a checkmark (✓) rather than an X

6.3.2.3.1 Widely Represented Indicators Across Assessments

A set of seven indicators demonstrates high coverage across the different assessment initiatives reviewed (scores of 7–9), including alignment with SDG monitoring priorities and partner requests. These include highest level of education, school ownership, school location, school size, number of teachers and staff, access to internet, and school facilities, capturing core institutional characteristics such as school type, enrollment size, and basic infrastructure and resource availability.

These constructs are consistently represented across both IEA and regional assessments and form the core of school-level contextual information. Their strong coverage reflects broad international consensus on the importance of school resources and structural conditions for supporting learning outcomes.

Several of these indicators are also relevant for SDG monitoring, with their role and contribution varying across indicators.

6.3.2.3.2 Moderate Coverage and Context-Sensitive School Environment Constructs

A second group of 17 indicators (scores of 4–6) shows moderate but consistent representation across multiple assessments. These include gender, age, years of service and experience in education, school type or structure, socioeconomic backgrounds of students, enrollment and number of classes in the target grade, language of test or instruction, school library and its use, availability of textbooks, provision of additional support or tutoring in reading or mathematics, digital learning resources, school year structure, days the school was closed, parental support, personal safety and bullying, student behavior and respect, and teacher absenteeism and availability. Taking together, these indicators capture school composition, principal characteristics, instructional organization, resource availability, and broader learning environment conditions.

These constructs extend beyond the basic structural characteristics covered in the 'widely represented' group and describe additional aspects of how schools are organized and operate as learning environments. Approximately one third of these indicators are also relevant for SDG monitoring.

6.3.2.3.3 Expert-Retained Low-Coverage Indicators

A set of indicators with limited cross-study coverage (scores ≤ 3) was retained for further analysis via expert override. These include professional development, school educational resources, school human resources, school year structure, shifts, instructional and curricular leadership and staff cohesion, student engagement and peer culture, school rules and enforcement, policies on grade retention and academic support, and nutrition or meals provided by the school.

Retention was based on their conceptual relevance and prior use in major assessments (including TIMSS/PIRLS and LaNA), indicating feasibility in diverse contexts. Collectively, these indicators capture organizational, resource-related, and policy dimensions of school functioning—such as staffing and time allocation, leadership and cohesion, rules and enforcement, and support provisions—which are pertinent in LMIC settings. Specific

considerations included the relevance of principal professional development, the use of shifts as a proxy for available instructional time and capacity, and the role of grade retention and support policies for managing learning progression.

6.3.2.3.4 Indicators Beyond the TIMSS/PIRLS Framework

Altogether, 12 retained indicators were not covered in the TIMSS/PIRLS school questionnaires. Most belong to the low-coverage indicators retained through expert judgment (7 of 12), although several are represented in the LaNA questionnaires. In addition, teacher job satisfaction is indirectly captured through principal reports in the absence of a dedicated teacher questionnaire and may therefore be considered an optional component when teacher data are not collected. Across the regional assessments, coverage is highest for principal background characteristics and school ownership, followed by school type and the provision of additional support.

6.3.3 Alignment with SDG Indicators and Identified Gaps

End-of-primary ILSA context questionnaires show in general a broad coverage but not always a direct alignment with the SDG indicators defined in the GAML framework. In many cases, available items proxy parts of an indicator (e.g., home background or learning environment) or rely on supplementary instruments such as parent questionnaires. Focusing on this project on student questionnaires only, several of these constructs on student level are not captured in IEA studies but may be in more detail included in non-IEA assessments. SDG indicators 4.4.1 (ICT skills) and 4.4.2 (digital literacy proficiency) are not directly measured at the end of primary level in any of the questionnaires investigated.

6.3.3.1 Coverage by SDG Indicators

The mapping of SDG indicators to questionnaire instruments presented in Table 6-12 provides an overview of this coverage. The table shows the questionnaire level, the SDG indicators judged as being measurable by ILSA context questionnaires according to the GAML framework (UIS, 2025), the coverage (with ‘xx’ indicating coverage via IEA TIMSS/PIRLS/LaNA and ‘x’ only covered in non-IEA assessments), an UIS rating conducted in April 2026, and additional information provided in a comment column.

In general, the table shows that IEA assessments cover SDG indicators at teacher and school level, while the coverage is lower at the student level. The table also shows that, while ILSAs provide a broad empirical basis for SDG-related measurement, the operationalizations do not always meet the full requirements of the indicators, as can be seen when comparing the full list of SDG indicators listed as in general suitable for ILSA measuring in combination with UIS rating which yield a mix of direct, weak, and non-alignments to the for this project finally selected and adapted questions.

At the student level, key background characteristics relevant for SDG indicators—such as gender, language, and migration status—are consistently captured across instruments. However, SDG indicators related to early learning conditions and family background are more unevenly represented. Only, information about the language of instruction and home language used (SDG 4.5.2) is robustly captured across IEA and non-IEA assessments while indicators linked to early childhood experiences and parental background are more frequently included in non-IEA assessments. In IEA assessments, these constructs are typically collected through parent or home questionnaires, reflecting the assumption that such information cannot be reliably reported by students. This particularly affects indicators such as participation in pre-

primary education (4.2.2 and 4.2.4), aspects of the home learning environment (4.2.3), and parental educational attainment (4.4.3), which are therefore not systematically included in IEA student questionnaires. In contrast, several non-IEA assessments incorporate elements such as parental education, pre-primary participation, or more comprehensive aspects of the home learning environment (SDG 4.2.3) directly within the student questionnaire, demonstrating the feasibility of simplified, student-reported proxy measures. As a result, the availability and comparability of these indicators depend both on instrument design choices and on whether supplementary questionnaires are administered.

At the school and teacher levels, SDG indicators are consistently represented across both IEA and non-IEA assessments, with two notable nuances. First, SDG 4.c.7 (in-service training over the past 12 months) is only explicitly captured in SEA-PLM, while most other assessments use longer reference periods. Second, additional items on initial teacher education and professional teaching qualifications which are more prevalent in non-IEA assessments can support a more comprehensive measurement of SDG 4.c.1 (teacher qualifications).

Table 6-12: SDG Indicators Measurable via ILSA Context Questionnaires and Coverage in R&D Maintained Indicators

Level	Indicator ID	SDG Description	Indicator	(at least partially) covered	UIS Assessment	Comments
Student	4.1.5	Children over-age for grade	Grade repetition	x	Not aligned	Non-IEA Assessments only. TIMSS/PIRLS home questionnaires cover 'age of child when beginning first grade' which is needed as well
Student	4.2.1	Developmentally on track (under 5s)				Home questionnaires
Student	4.2.2	Organized learning participation	Participation in formal pre-primary education	x	Directly aligned	Non-IEA Assessments only, IEA parent questionnaires
Student	4.2.3	Stimulating home environments	Number of books/ Parental involvement in learning/ Parents' literacy/	x	Directly aligned/ Directly aligned/ Weak alignment	Parental reading habits included in PIRLS parent questionnaire; Non-IEA assessments include more direct information on parent's involvement

Level	Indicator ID	SDG Description	Indicator	(at least partially) covered	UIS Assessment	Comments
Student	4.2.4	ECE enrollment rate	Participation in formal pre-primary education	x	Directly aligned	Non-IEA Assessments only, IEA parent questionnaires
Student	4.4.1	ICT skills by type				Only indirectly via ICT related questions in teacher questionnaires
Student	4.4.2	Digital literacy proficiency				
Student	4.4.3	Educational attainment rates	Parents' Highest formal education	x	Weak alignment	SEA-PLM only, otherwise included in parent questionnaires
Student	4.5.2	Home language instruction	Language experience	xx	Directly aligned	
School	4.a.1	Basic school services/infrastructure	School facilities (incl. Toilets SDG)/Internet (contr.)/Digital learning resources (contr.)/School- or Classroom Library and use (contr.)	xx	N/R*	
School	4.a.2	Experience of bullying	SOS_Personal Safety / BULLYING (incl. Teachers)	xx	Directly aligned	
Teacher	4.c.1	Teacher qualifications	Highest level of education (Teacher)/Initial Teacher Education (duration/quality/Experiences)/ Professional Teaching Qualification	xx	Directly aligned	Non-IEA assessment add to Initial Teacher Education and Professional Teaching Qualification
Teacher	4.c.2	Pupil-trained teacher ratio	Professional development (duration, quality)	xx	Weak alignment	Needs enrollment data as denominator
Teacher	4.c.4	Pupil-Qualified Teacher Ratio	Highest level of education (Teacher)/ Professional Teaching Qualification/ Area(s) of Post-secondary Study / Teacher	xx	Directly aligned	Needs enrollment data as denominator

Level	Indicator ID	SDG Description	Indicator	(at least partially) covered	UIS Assessment	Comments
Teacher	4.c.7	In-service training (12m)	Professional development (duration, quality)	x	Weak alignment	Only SEA-PLM adhering to 12m timeframe

xx - Included in IEA and Non-IEA assessments at the respective level (Student, Teacher, School)

x - Included only in Non-IEA assessments at the respective level

* - Indicator was not reviewed by UIS due to an incorrect indicator name submitted during reporting

6.3.3.2 Disaggregation Variables and Prerequisites

In addition to the direct measurement of SDG indicators, the context questionnaires capture a set of variables that function as essential disaggregation categories or as prerequisites for the construction of SDG indicators. As shown in Table 6-13, variables such as gender, migration background, school location, and language are widely available across both IEA and non-IEA assessments and support the disaggregation requirements of the SDG monitoring framework. Similarly, information on household possessions and services, as well as parental education, enables the construction of socioeconomic indices that are frequently used for equity analyses and to construct a wealth index which then in turn use a disaggregation variable for different SDG indicators. School-level data on enrollment and class size provides the necessary inputs for computing ratio-based indicators, particularly those related to teacher supply.

Some differences exist in the types of socioeconomic indicators included. Variables related to housing quality, such as the type or material of the dwelling, are only present in non-IEA assessments, and items included in household possessions and services also differ. This pattern reflects the stronger orientation of regional assessments toward LMIC contexts.

6.3.3.3 Patterns across Tables

As discussed above, differences between IEA and non-IEA assessments are most pronounced at the student level and largely reflect instrument design choices, particularly the reliance on parent questionnaires in IEA studies versus more direct inclusion of constructs in non-IEA student questionnaires (even if parent questionnaires are administered as well), affecting the consistency of SDG indicator coverage across contexts.

A common limitation across all assessments concerns indicators that require ratio-based calculations, notably SDG indicators 4.c.2 and 4.c.4. While teacher questionnaires capture key components such as qualifications and training, the computation of these indicators requires additional information on student enrollment that must be linked from school-level data, highlighting the need for integrated data structures across questionnaires. See Table 6-13 for information on SDG indicators that are useful as disaggregation measures or metric prerequisites.

Table 6-13: SDG Indicators useful as Disaggregation measures or Metric Prerequisites

Level	SDG Indicator(s)	Function	Indicator	(at least partially) covered
Student	4.1.1, 4.5.1	Disaggregation variable	Gender	xx

Level	SDG Indicator(s)	Function	Indicator	(at least partially) covered
Student	4.1.5	Metric Prerequisite	Age	xx
Student	4.1.5, 4.5.2	Disaggregation variable	Immigration Status	xx
Student	4.1.1, 4.5, and others	Useful for creating wealth index as disaggregation	Household possessions and services	xx
Student	4.1.1, 4.5, and others	Useful for creating wealth index as disaggregation	Type, material of house	x
Teacher	4.c.1, 4.c.7	Disaggregation variable	Teacher Gender	xx
School	4.1.1, 4.5.1	Disaggregation variable	School location (urban/rural)	xx
School	4.1.1, 4.5.1	Disaggregation variable	SES Backgrounds of students	xx
School	4.c.2	Metric Prerequisite	School size (total enrollment of boys and girls and at target grade)	xx
School	4.c.2	Metric Prerequisite	Enrollment & number of classes in target grade	xx
School	4.5.2	Disaggregation variable	Language of test/ language of instruction	xx
School	4.c.1	Metric Prerequisite	Number of teachers/ Staff	xx

xx - included in IEA and Non-IEA assessments at the respective level (Student, Teacher, School)

x - included only in Non-IEA assessments at the respective level

6.4 Operationalization of Indicators

The operationalization stage resulted in a structured mapping of all retained indicators to one or two questionnaire items. As described in 5.3.1.2, the indicators themselves were generated through a systematic comparison and grouping of conceptually similar questionnaire items across the included assessments, rather than by aligning pre-existing construct labels. This mapping represents the final outcome of the item identification, selection, and adaptation process described in Section 5.4 and defines the set of items carried forward for subsequent psychometric and qualitative evaluation.

For each indicator, items were selected from the compiled cross-assessment item pool. This pool was derived from the prior indicator creation step, which matched items across international and regional (LMIC) assessments to common indicators and produced, for each retained indicator, a list of all associated questions. The selected items were then reviewed in

detail with regard to their alignment with the intended construct, respondent suitability, and feasibility of administration. Where necessary, items were revised by the project's research team to improve clarity, reduce cognitive burden, or enhance contextual relevance for LMIC settings. In some cases, additional items were combined or newly introduced by the research team to ensure adequate coverage of the indicator.

The proposed questionnaire items represent a practical operationalization of the retained indicators rather than a comprehensive measurement of every aspect of the underlying construct. In some cases, particularly for student questionnaires, operationalizations were intentionally simplified to ensure age appropriateness and reduce respondent burden. Consequently, although individual items were selected to capture the core aspect of an indicator, the proposed set of questions may not always fully represent all dimensions of the constructs identified during the questionnaire-matching step. The resulting item set listed below reflects the application of the hierarchical source strategy. Wherever suitable items from IEA studies (TIMSS, PIRLS, LaNA) were available, these were retained as the primary operationalization to support comparability. Where multiple candidate items existed across assessments, the simplest effective formulation was selected. Regional assessments contributed complementary or alternative items in cases where existing IEA items were not sufficiently applicable or did not capture context-specific aspects of the construct. The starting point for item formulation was a common framework rooted in TIMSS and PIRLS; wording and response categories were adjusted where necessary to align with this style. In some cases, these adjustments resulted in substantial modifications of the originally selected items. To document the degree of similarity to the source items, an adaptation classification was developed, which is described in a later section.

The operationalization further reflects consistent application of the predefined design principles for questionnaire development, including simplified wording, limited response categories, and avoidance of complex routing—particularly for student instruments. Efforts were made to preserve conceptual alignment with the original constructs across all item versions. The finalized items are presented in analytical form in Section 6.4.1. For readability, all items suggested for the different questionnaires are provided in a questionnaire layout in Appendix 8.1 (A1 - Student) 8.2 (A2 – Teacher), and 8.3 (A3 – School).

This section documents the finalized item pool as it emerged from the internal review and adaptation process. Empirical characteristics and qualitative assessments are described in the subsequent chapters.

6.4.1 Student-Level Operationalization

At the student level, the operationalization resulted in a set of 27 indicators, represented through a total of 36 questionnaire items (see

Table 6-14), corresponding to an average of approximately 1.3 items per indicator. While most indicators were measured using a single item, a subset of 9 indicators was operationalized using two items (A/B). These multi-item operationalizations were primarily used for more complex or multi-dimensional constructs, particularly in areas such as language experience, absenteeism, household composition, health, and nutrition/meals. In addition, multi-item operationalizations were used to more comprehensively capture students' learning-related experiences, specifically targeting reading and mathematics contexts. Given the need to keep the student questionnaire age-appropriate and concise, some indicators are represented by items that capture only their central dimension rather than the full underlying construct. For

example, the proposed student birthplace item captures one aspect of immigration status but does not provide a complete classification of migration background, which would require additional information such as parental or guardian birthplace.

Table 6-14: Student Level Mapping

# 22	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
1	Gender	A	Which of these describes you?	Girl; Boy; <Other>	LaNA	Moderate
2	Age	A	How old are you?	---	LaNA	Minimal
3	Language experience	A	How often do you speak the <language of the test> at home?	Always; Almost always; Sometimes; Never	LaNA	Minimal
3	Language experience	B	What language do you speak at home most of the time?	<language of test>; <language 2>; <language 3>; <Other> (Note: If...	SEA-PLM	Minimal
4	Immigration Status	A	Where were you born?	In this same...; In another...; In another country; I don't know	TIMSS	Substantial
5	Parents' literacy	A	How many adults in your home can read or write?	No adults at home...; One adult at home...; More than one...; I don't know	PASEC	Substantial
6	Parents' Highest formal education	A	What is the highest level of education that your father (or...	Mother or female...; Father or male...	SEA-PLM	Moderate
7	Participation in formal pre-primary education	A	Did you go to a preschool or kindergarten before starting...?	Yes, for 2 years...; Yes, for between...; Yes, for less...; No; I don't know	SEA-PLM	Moderate
8	Grade repetition	A	Have you ever repeated a grade?	Yes; No	SEA-PLM	Minimal
9	Absenteeism	A	About how often are you absent from school?	Never/almost never; About once every...; About once per...; About once every...; About once a week...	LaNA	Moderate
9	Absenteeism	B	About how often do you arrive late to school?	Never/almost never; About once every...; About once per...; About once every...; About once a week...	ERCE	Substantial

²² The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

# 22	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
10	Household composition (number of people living at home)	A	How many people currently live in your home? (Including...	1-3; 4-6; 7-9; 10 or more	LaNA	Moderate
10	Household composition (number of people living at home)	B	How many siblings (brothers and sisters) do you have?	None; 1; 2; 3; 4; More than 4	SEA-PLM	Substantial
11	Number of books in the home	A	How many books are there in your home? (Do not count...	No books; 1-10; 11-20; 21-30; 31-50; 51-100; More than 100	LaNA	Moderate
12	Household possessions and services	A	Do you have any of these things at your home? a) Running...	Yes; No	LaNA/ TIMSS	Substantial
13	Type, material of house	A	What are the walls of your home mostly made of?	Cardboard; Grass; Mud and sticks; Stones; Metal sheets; Wood; Cut stones or...; Other	SACMEQ	Substantial
14	Parental involvement in learning (incl. help with schoolwork/ Tuition)	A	How often does someone at home help you with your homework?	I do not get any...; Never; Sometimes; Most of the time	PASEC	Substantial
15	Assignment of Homework	A	About how often do you get homework? a) In mathematics b)...	I do not get any...; 1-2x/month; 1-2x/week; Most days of the...	SACMEQ	Moderate
16	Enjoyment of reading/ mathematics	A	How much do you agree with these statements about reading?...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	PIRLS	Moderate
16	Enjoyment of reading/ mathematics	B	How much do you agree with these statements about learning...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	TIMSS	Moderate
17	Confidence in reading/ mathematics	A	How much do you agree with these statements about reading?...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	PIRLS	Moderate
17	Confidence in reading/ mathematics	B	How much do you agree with these statements about...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	TIMSS	Moderate
18	Nutrition, meals	A	Which meals do you eat on a normal school day? a) Breakfast...	Yes; No	SEA-PLM	Substantial
18	Nutrition, meals	B	How often do you have trouble paying attention in school...	Daily; Almost daily; Sometimes; Never	LaNA	Substantial

# 22	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
19	Physical and mental health (incl. Anxiety, Depression)	A	How often do you have trouble paying attention in school...	Daily; Almost daily; Sometimes; Never	LaNA	Substantial
19	Physical and mental health (incl. Anxiety, Depression)	B	Do you agree with these statements about your health: a) I...	Yes; No	PISA_D	Substantial
20	Well being at school	A	How much do you agree with these statements about your...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	LaNA	Minimal
21	Classroom climate	A	How often do these things happen in your class? a) Students...	Every or almost...; 1-2x/week; 1-2x/month; Never/almost never	PIRLS	Moderate
22	Experience of bullying/ harassment	A	During this year, how often have other students from your...	At least once a...; 1-2x/month; A few times; Never	PIRLS	Minimal
23	Access to books from community/school/class library	A	How often do you borrow books (including ebooks) from your...?	I do not have...; At least once a...; 1-2x/month; A few times a year; Never/almost never	PIRLS	Minimal
24	House chores/ Work	A	How often do you do these things outside of school? a) Do...	Daily; Almost daily; Sometimes; Never	SEA-PLM	Moderate
25	Commuting to school	A	How long does it take you to get to school each day?	Less than 30...; 30 minutes up to...; 1 hour to 2 hours; 2 hours or more	LaNA	Minimal
26	Language of class instruction	A	How often does your reading/ language teacher use a...?	Never; Sometimes; Often	SACMEQ	Substantial
26	Language of class instruction	B	Which other language does your language teacher use the...?	None, my teacher...; <language 1>; <language 2>; <language 3>; <Other>	NEW	Substantial
27	Perception of Instruction	A	How much do you agree with these statements about your...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	LaNA	Substantial
27	Perception of Instruction	B	How much do you agree with these statements about your...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	LaNA	Substantial

X - indicator available in student questionnaire

(X) - indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

6.4.1.1 Coverage Patterns

The coverage pattern at the student level reflects the indicators selected in the preceding selection process, with a strong emphasis on background characteristics and learning-related

experiences relevant for LMIC contexts, including demographic variables (e.g., age, gender), language exposure, early childhood education, and aspects of well-being and health. These constructs are generally amenable to direct self-reporting and could therefore be implemented using relatively simple item formats.

Most indicators were operationalized using single-item measures, reflecting the suitability criteria applied for this respondent group, particularly regarding minimizing cognitive burden and ensuring clarity for younger learners. Multi-item operationalizations were used for constructs requiring additional differentiation, particularly where complementary aspects of a construct needed to be captured. These include language experience, absenteeism, household composition, health, nutrition, and language of instruction. A second group of indicators, each operationalized with two items, comprises students' subject-specific experiences in reading and mathematics, covering enjoyment, confidence, and perception of instruction.

6.4.1.2 Source Patterns

In terms of their questionnaire sources, the suggested student questionnaire items were classified into three categories: items derived exclusively from IEA instruments, items derived exclusively from non-IEA (regional) assessments, and non-IEA frameworks, and newly developed items (NEW).

Across the 36 items, a total of 21 items were derived exclusively from IEA instruments (TIMSS, PIRLS, and LaNA), primarily covering constructs with well-established international definitions such as demographic characteristics and core learning-related experiences. In addition, 14 items originated exclusively from non-IEA regional assessments, contributing mainly to domains such as language use, household characteristics, and parental background in LMIC settings. One substantially adapted item was based on PISA-D, which uniquely captures a broad range of health issues. While some of these constructs, such as parental background and preschool experiences, are also included in TIMSS and PIRLS frameworks, they are typically collected via parent questionnaires. The present item set demonstrates that these constructs can, in adapted form, also be captured through student reports, which is particularly relevant in contexts where parent questionnaires are not administered. Finally, 1 item was newly developed or substantially redesigned (NEW) based on professional judgment, taking the available sources into account.

Overall, this distribution reflects a construct-driven integration strategy, balancing international comparability with contextual relevance and feasibility for student respondents in LMIC settings.

6.4.1.3 Adaptation Patterns

The adaptation pattern shows a diverse range of modifications across the 36 items. In total, 8 items (22%) required minimal adaptation, 12 items (33%) required moderate adaptation, and 16 items (44%) required substantial adaptation. While minimal adaptations were typically limited to alignment of wording or response formats (e.g., age and commuting time), moderate and substantial adaptations were applied to ensure feasibility for Grade 4–6 students in the LMICs context and to cover indicators more comprehensively. Substantial adaptations were primarily

applied to complex or context-sensitive indicators such as immigration status, parents' literacy, household possessions, and homework involvement.

6.4.2 Teacher-Level Operationalization

At the teacher level, operationalization yielded 27 indicators represented by 36 questionnaire items (see Table 6-15). Most indicators were measured using a single item, while 9 indicators were operationalized using two items (A/B), primarily for more complex constructs related to instructional practices, professional experience, and teacher qualifications.

Table 6-15: Teacher Level Mapping

# ²³	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
1	Teacher Gender	A	Which of these describes you?	Female; Male; <Other>	TIMSS	Moderate
2	Teacher Age	A	How old are you?	<25; 25–29; 30–39; 40–49; 50–59; 60 or more	TIMSS	Minimal
3	Years of Teaching	A	By the end of this school year, how many years will you...?	----- years...	TIMSS	Minimal
3	Years of Teaching	B	By the end of this school year, how many years will you...?	----- years...	TIMSS/PISA-D	Substantial
4	Highest level of education	A	What is the highest level of formal education you have...?	Did not complete...; <Upper secondary...; <Post-secondary...; <Short-cycle...; <Bachelor's or...; <Master's or...; <Doctoral or...	TIMSS	Minimal
5	Professional Teaching Qualification	A	Do you have a formal teaching qualification that qualifies...	Yes; No	ERCE	Moderate

²³ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

# ²³	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
6	Subject-Matter Assignments (general /in sel. class)	A	Which of the following two subjects do you teach in the...	Yes; No	ERCE	Moderate
7	Grades/ Classes being taught	A	Which <grade(s) or class level(s)> do you currently teach...?	<Grade or class...; <Grade or class...; <Grade or class...; <Grade or class...; <Grade or class...; <Grade or class...; <Other>	ERCE/PISA-D	Substantial
8	Initial Teacher Education (duration/quality/Experiences)	A	Did you complete any pre-service teacher training? (This...	No; Yes, for up to 1...; Yes, for more...; Yes, for more...; Yes, for more...	PISA-D	Substantial
9	Area(s) of Post-secondary Study / Teacher	A	During your <post-secondary> education, what was your major...	Yes; No	TIMSS	Moderate
9	Area(s) of Post-secondary Study / Teacher	B	If your major or main area of study was education, did you...	Yes; No	TIMSS	Minimal
10	Professional development (duration, quality)	A	In the past two years, have you participated in formal...	Participated: Yes; Participated: No; Importance: Not...; Importance...; Importance: Very...	TIMSS/PIRLS	Substantial
10	Professional development (duration, quality)	B	If you answered yes to any of the activities above, to what...?	A lot; Some; A little; Not at all	ERCE	Moderate
11	Confidence in teaching (methods)	A	How confident are you in doing each of these things when...	V. high; High; Medium; Low	ERCE/TIMSS	Moderate

# ²³	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
12	Job satisfaction (incl. General working conditions)	A	How often do you have these feelings about being a teacher?...	Very often; Often; Sometimes; Never	TIMSS	Minimal
13	Teacher Absence	A	During the last month, did any of the following situations...	Yes;; No	PISA-D	Substantial
14	Type of contract with school/ Position	A	How long is your employment contract with this school?	For an indefinite...; For a term of one...; For a term of...; Other	ERCE	Substantial
15	Additional paid jobs	A	In addition to your regular teaching at this school, how...?	None; Fewer than 10...; 10 to 19 hours...; 20 to 30 hours...; More than 30...	PISA-D	Moderate
16	Instructional Resources and Classroom Infrastructure Inventory	A	Which of the following items do you have in your classroom?...	Yes; No	SEA-PLM	Substantial
16	Instructional Resources and Classroom Infrastructure Inventory	B	How does your school make textbooks available to students...	There are no...; One textbook is...; One textbook is...; Most students...; Each student has...	PISA-D	Substantial
17	School- or Classroom library and use	A	How many books with different titles are in your classroom...?	My classroom does...; Fewer than 25; 26 to 50; 51 to 100; More than 100	PIRLS	Substantial
17	School- or Classroom library and use	B	Can the students borrow books from the classroom library or...?	Yes; No	PIRLS	Minimal
18	EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)	A	How would you characterize each of the following within...	V. high; High; Medium; Low; Very Low	TIMSS	Minimal
19	EAS_Parental Support	A	How would you characterize each of the following within...	V. high; High; Medium; Low; Very Low	TIMSS	Minimal

# ²³	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
20	SOS_Personal Safety (including bullying)	A	How often have these things happened at your school in the...	At least once a...; 1–2x/month; A few times; Never	ERCE	Substantial
20	SOS_Personal Safety (including bullying)	B	How much do you agree or disagree with the following...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	TIMSS/PIRLS	Moderate
21	SOS_Student Behavior & Respect	A	How much do you agree or disagree with the following...	Agr. a lot; Agr. a little; Disagr. a little; Disagr. a lot	TIMSS	Substantial
21	SOS_Student Behavior & Respect	B	How often do the following situations occur in the...	Very often; Often; Sometimes; Never	ERCE	Moderate
22	External Supervision & Inspection	A	Since the beginning of this school year, how often have the...	At least once a...; 1–2x/month; A few times a year; Never	PASEC	Substantial
23	Class size/ number of students on day of test	A	How many students are in this class? a) How many students...?	Male students...; Female students...	TIMSS	Substantial
24	Teaching conditions/ Teaching Constraints	A	In your view, to what extent do the following limit how you...	Not at all; Some; A lot	TIMSS	Minimal
24	Teaching conditions/ Teaching Constraints	B	To what extent do the following limit how you teach this...	Not at all; Some; A lot	NEW	Substantial
25	Language of instruction	A	What languages do you use when teaching your students in...?	Only <language of...; Mostly <language...; Sometimes...; Mostly students'...; Always students'...	PISA-D	Minimal
26	Instructional time (Language/ Reading or mathematics)	A	In a typical week, how much time do you spend teaching...?	----- minutes...	TIMSS	Minimal

# ²³	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
26	Instructional time (Language/ Reading or mathematics)	B	In a typical week, how much time do you spend on <language...?	----- minutes...	PIRLS	Minimal
27	Instructional activities (general, Language/ Reading, Math)	A	How often do you do the following in teaching this class?...	Every or almost...; About half the...; Some lessons; Never	TIMSS	Moderate

X = indicator available in student questionnaire

(X)=indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

6.4.2.1 Coverage Patterns

The coverage pattern at the teacher level reflects the indicators selected in the preceding selection process, with a strong emphasis on teacher background, professional qualifications, instructional practices, and classroom-related experiences. Based on the study classification, a substantial share of items originates from large-scale international assessments focusing on teaching practices and professional characteristics. These constructs sometimes require slightly more complex formulations compared to the student level, but questions could mainly be operationalized relying on simple response formats and avoiding filter questions.

Most indicators were operationalized using single-item measures, reflecting the continued emphasis on feasibility and clarity. Multi-item operationalizations were used selectively for constructs requiring additional differentiation, particularly where multiple dimensions of teaching practice, professional experience, or subject specific questions needed to be captured.

6.4.2.2 Source Patterns

The suggested teacher questionnaire items fell into four source categories: IEA-only, non-IEA-only, mixed-source, and newly developed (NEW).

Across the 36 items, the source pattern shows a clear dominance of IEA-based items, with 19 items derived exclusively from IEA instruments (TIMSS, PIRLS). These primarily cover constructs related to teacher background, qualifications, and instructional practices. A substantial share of 14 items originated exclusively from non-IEA regional assessments (e.g., ERCE, PISA-D, SEA-PLM, PASEC), contributing mainly to context-specific aspects of teaching conditions and professional environments.

In addition, 2 items represent mixed-source adaptations, combining IEA and non-IEA elements, and 1 item was newly developed (NEW) to address a construct not sufficiently covered in existing frameworks.

6.4.2.3 Adaptation Patterns

The adaptation pattern shows a range of modifications across the 36 items. Twelve items (33%) required only minimal alignment, ten items (28%) required moderate adaptation, and fourteen items (39%) required substantial modification.

Minimal adaptations were typically limited to alignment of wording or response formats, while moderate adaptations were applied to ensure consistency in terminology and response

structures. Substantial adaptations were used for constructs requiring contextual adjustment, particularly where items were drawn from non-IEA sources or needed to be harmonized with the overall questionnaire framework.

Compared to the student level, adaptation needs were generally less driven by cognitive constraints and more by terminological alignment and the need to maintain broad contextual relevance, particularly for professional and instructional constructs.

6.4.3 School-Level Operationalization

At the school level, operationalization yielded 36 indicators across 41 items (see Table 6-16).

Most indicators were measured using a single item, while 5 indicators were operationalized using two items (A/B). These were used for constructs requiring additional differentiation, primarily where multiple aspects of school conditions or organizational structures needed to be captured.

Table 6-16: School Level Mapping

# ²⁴	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
1	Gender	A	Which of these describes you?	Female; Male; <Other>	SEA-PLM	Moderate
2	Age	A	How old are you?	<25; 25–29; 30–39; 40–49; 50–59; 60 or more	SEA-PLM	Minimal
3	Years of service, experience in education	A	By the end of this school year, how many years will you...?	___ years	TIMSS/PIRLS	Minimal
3	Years of service, experience in education	B	By the end of this school year, how many years will you...?	___ years	TIMSS/PIRLS	Minimal
4	Highest level of education	A	What is the highest level of formal education you have...?	Did not complete...; <Upper secondary...; <Post-secondary...; <Short-cycle...; <Bachelor's or...; <Master's or...; <Doctoral or...	TIMSS/PIRLS	Moderate
5	Professional development	A	Have you completed any formal professional training in...?	Yes; No	ERCE/PASEC	Substantial

²⁴ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

# ²⁴	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
6	School ownership (private/public/religious)	A	Which of these options best describes your school?	Public: managed...; Private/...	SEA-PLM	Substantial
6	School ownership (private/public/religious)	B	What kind of organization manages your school?	A church or other...; A not-for-profit...; A for-profit...; The government; Other	PISA-D	Substantial
7	School type/structure (primary/secondary/language)	A	Which of the following levels of education does your school...?	<Early childhood...; <Primary...; <Lower secondary...; <Upper secondary...	PISA-D	Minimal
8	School location (urban/rural)	A	How many people live in the city, town, or area where your...?	More than 500,000...; 100,001–500,000...; 50,001–100,000...; 30,001–50,000...; 15,001–30,000...; 3,001–15,000...; 3,000 people or...	TIMSS/PIRLS	Minimal
8	School location (urban/rural)	B	Which best describes the immediate area in which your...?	Urban—Densely...; Suburban—On...; Medium size city...; Small town or...; Remote rural	LaNA/TIMSS	Minimal
9	SES Backgrounds of students	A	Approximately what percentage of students in your school...	0–10%; 11–25%; 26–50%; 51–75%; More than 75%	PIRLS/TIMSS	Moderate
10	School size (total enrollment of boys and girls and at target grade)	A	What is the total enrollment of students in your school as...	___ boys; ___ girls	TIMSS	Substantial
11	Enrollment & number of classes in target grade	A	What is the total enrollment of <target grade> students in...	___ boys; ___ girls	TIMSS/PIRLS	Substantial
12	Language of test/ language of instruction	A	Approximately what percentage of students in your school...?	More than 90%; 76–90%; 51–75%; 26–50%; 25% or less	TIMSS	Minimal

# ²⁴	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
13	Number of teachers/ Staff	A	How many teachers are at your school? (A full-time teacher...	___ full time...; ___ part time...; ___ part time...; ___ part time...	LaNA	Substantial
14	School Educational Resources except books (and affection by shortage)	A	How much is your school's capacity to provide instruction...	Not at all; A little; Some; A lot	LaNA	Moderate
15	School human resources (and affected by shortage)	A	How much is your school's capacity to provide instruction...	Not at all; A little; Some; A lot	LaNA	Moderate
16	School Library (equipment and use)	A	Approximately, how many books does your school library have?	My school does...; 50 books or fewer; 51 to 100 books; 101 to 200 books; 201 to 500 books; 501 to 1000 books; 1001 books or more	PIRLS/SEA-PLM	Substantial
17	Internet	A	Does your school have the following resources? a) Reliable...	Yes; No	TIMSS/LaNA	Moderate
18	School facilities (incl. Toilets SDG)	A	Does your school have the following resources? a)...	Yes; No	ERCE	Substantial
19	Classroom facilities	A	Does your school have the following classroom resources? a)...	Yes; No	LaNA	Substantial
20	Availability of textbooks	A	How does your school make textbooks available to students...	There are no...; One textbook is...; One textbook is...; Most students...; Each student has...	SEA-PLM	Moderate
20	Availability of textbooks	B	How much is your school's capacity to provide instruction...	Not at all; A little; Some; A lot	LaNA	Moderate
21	Provision of additional support in reading or math	A	Does your school provide extra classes or tutoring for...	Yes, at no...; Yes, for a fee; No	AMPL/SEA-PLM	Moderate

# ²⁴	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
22	Digital learning resources	A	How many digital devices does your school have for use by...?	___ devices	PIRLS	Substantial
23	School Year Structure (beginning/end of school year, operational days)	A	For the <target grade> students in your school: How many...?	___ days	LaNA	Minimal
23	School Year Structure (beginning/end of school year, operational days)	B	In one calendar week, how many days is the school open for...?	6 days; 5 1/2 days; 5 days; 4 1/2 days; 4 days; Other	LaNA	Minimal
24	School Year Structure (instructional time per day)	A	What is the total instructional time, excluding breaks, on...?	___ minutes	LaNA	Minimal
25	Shifts	A	Does your school have double shifts or double sessions?...	Yes; No	PASEC	Minimal
26	Days school closed (Covid, strikes, etc.)	A	Within the last 12 month, about how many days was your...?	_____ days	LaNA	Minimal
27	Teacher job satisfaction	A	How would you characterize each of the following within...	V. high; High; Medium; Low; V. low	LaNA	Minimal
28	EAS_Instructional and Curricular Leadership and Staff Cohesion	A	How would you characterize each of the following within...	V. high; High; Medium; Low; V. low	LaNA	Minimal
29	EAS_Parental Support	A	How would you characterize each of the following within...	V. high; High; Medium; Low; V. low	LaNA	Minimal
30	EAS_Student Engagement and Peer Culture	A	How would you characterize each of the following within...	V. high; High; Medium; Low; V. low	LaNA	Minimal
31	SOS_Personal Safety / BULLYING (incl. Teachers)	A	To what degree is each of the following a problem among...	Not a problem; Minor problem; Moderate problem; Serious problem	LaNA	Moderate
32	SOS_School rules and enforcement	A	To what degree is each of the following a problem among...	Not a problem; Minor problem; Moderate problem; Serious problem	LaNA	Minimal

# ²⁴	INDICATOR	Item	Suggested Item Stem	Suggested Item Options	Source	Adaptation
33	SOS_Student Behavior & Respect	A	To what degree is each of the following a problem among...	Not a problem; Minor problem; Moderate problem; Serious problem	LaNA	Minimal
34	Policies on grade retention/ academic support & digital devices	A	What are your school's policies regarding <grade...	Yes; No	PISA	Minimal
35	Teacher Absenteeism, Teacher availability, Issues at school	A	To what degree is each of the following a problem among...	Not a problem; Minor problem; Moderate problem; Serious problem	LaNA	Moderate
36	Nutrition, meals provided by school	A	Does your school provide free meals for students? a)...	Yes, for all...; Yes, for some...; No	LaNA	Minimal

X - indicator available in student questionnaire

(X) - indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

Note: 3 indicators (AIDS related/ # of students and teachers died, Covid Pandemic, Questionnaire rating) were excluded from this project due to limited cross-comparability and relevance.

6.4.3.1 Coverage Patterns

The coverage pattern at the school level reflects the indicators selected in the preceding selection process, with an emphasis on school characteristics and organizational features of the learning environment. Indicators predominantly capture structural constructs such as school resources, infrastructure, staffing conditions, school climate, and organizational practices, which are well-defined and can be reported by school principals or administrators with relatively low ambiguity.

As a result, most indicators could be operationalized using single-item measures, reflecting both the clarity of the constructs and the feasibility of collecting this information in large-scale survey settings. Multi-item operationalizations were used only in cases where multiple components of school conditions needed to be captured to cover the respective indicator more comprehensively.

6.4.3.2 Source Patterns

The suggested school questionnaire items fell into three source categories: IEA-only, non-IEA-only, and mixed-source.

Across the 41 items, the source pattern shows a strong dominance of IEA-based items (29 items). A large majority of these items were taken from the LaNA 2023 administration, which itself builds on established TIMSS and PIRLS frameworks and demonstrates their applicability in LMIC contexts. These items primarily cover constructs related to school resources, infrastructure, staffing conditions, and school organization.

A little over one-third (11 items) originated from non-IEA regional assessments (SEA-PLM, ERCE, PASEC, AMPL, and PISA-D), contributing mainly to context-specific aspects of school

conditions in LMIC settings as well as personal background information about the principal. One item on the number of books available in the school library was adapted based on PIRLS and SEA-PLM question elements, representing a mixed-source adaptation.

6.4.3.3 Adaptation Patterns

The adaptation pattern shows a comparatively lower level of modification than observed for the student and teacher levels. About half of the items (20 items, 49%) required only minimal adaptation, while 11 items (27%) required moderate adaptation and 10 items (24%) required substantial adaptation.

Adaptations were primarily applied to ensure contextual relevance for LMIC settings, particularly for indicators related to infrastructure, resource availability, and school-level constraints. Many of the LaNA items, already administered in LMIC contexts and aligned with IEA conventions, were classified as requiring only minimal adaptation.

6.4.4 Cross-Level Patterns in Sources and Adaptation

6.4.4.1 Source Distribution Comparison

Across the three respondent levels, differences in source distributions are closely linked to the nature of the constructs being measured and their typical placement within international assessment frameworks.

At the student level, the item set includes a substantial share of constructs related to home background and student context, such as socioeconomic conditions, health, nutrition, and language use. In IEA studies such as TIMSS and PIRLS, many of these constructs are traditionally collected through parent questionnaires, reflecting the expectation that parents are better positioned to provide accurate information. In contrast, several regional assessments administered in LMIC contexts rely more extensively on student questionnaires to capture these aspects, even if parent questionnaires are administered. As a result, the student-level operationalization included a broader share of non-IEA item sources, which in turn often required additional adaptation to align with IEA structure and wording.

At the teacher level, the source pattern reflects a combination of established IEA constructs and additional indicators related to teacher background and classroom organization, which are particularly relevant in LMIC contexts. While core areas such as professional qualifications, teaching experience, and instructional practices are well covered by IEA frameworks, complementary items from regional assessments were required to capture aspects such as contractual conditions, teaching assignments, or classroom-level constraints.

At the school level, the source distribution shows the strongest reliance on IEA-based items, largely derived from the LaNA 2023 administration and its underlying TIMSS and PIRLS frameworks, which already demonstrated broader coverage of topics relevant in LMIC contexts. School-level constructs—such as infrastructure, staffing, and organizational practices—are comparatively stable and well established across international assessments. Consequently, the need for non-IEA source items is more limited and primarily restricted to context-specific indicators related to resource constraints or administrative conditions.

6.4.4.2 Adaptation Distribution Comparison

Differences in adaptation patterns across levels follow a similar logic and are closely linked to both respondent characteristics and construct complexity.

At the student level, a large proportion of items required moderate or substantial adaptation. This reflects the need to simplify wording, reduce cognitive load, and adjust response formats to ensure that questions are understandable for students in Grades 4–6. Adaptations were particularly pronounced for constructs related to home background, where items were reformulated to be answerable by students rather than parents.

At the teacher level, the proportion of moderate adaptations remains substantial, while the share of minimal adaptations increases.

At the school level, the extent of adaptation is lower, with nearly half of all items requiring only minor adjustments. Many items were taken directly from LaNA and closely follow established IEA formats, which have already been shown to function in LMIC contexts. Moderate and substantial adaptations are primarily limited to indicators capturing context-specific aspects of infrastructure, resource availability, and institutional constraints.

6.4.4.3 Cross-Level Summary

The observed patterns show a systematic alignment between construct type, respondent group, and operationalization strategy.

At the student level, the item set includes a higher share of context-dependent constructs, particularly related to home background, which are often captured via parent questionnaires in IEA studies. In this study, these constructs are measured through student reports, which is reflected in the broader integration of non-IEA items and the higher degree of adaptation.

At the teacher level, the item set combines established IEA constructs with additional indicators related to teacher background and classroom organization. This results in a moderate level of source diversification and adaptation.

At the school level, the item set relies predominantly on standardized IEA-based items, with comparatively limited use of non-IEA sources and lower levels of adaptation.

6.5 Empirical Characteristics of Proposed Indicators and Selected Questions

Following the content-based construction and selection of indicators and questions, an empirical screening was conducted to assess the analytical suitability of the retained variables within LMIC contexts. This step aimed to provide descriptive evidence on response behavior, distributional properties, and empirical relevance for learning outcomes, thereby complementing the conceptual and expert-based selection procedures.

The analyses were based exclusively on LMIC country samples across all included assessments, ensuring alignment with the study's LMICs focus. All statistics were aggregated on school level and computed using study-specific student weights, and achievement measures were harmonized across studies by aggregating plausible values where available or using single-score outcomes otherwise. Where indicators were measured through multiple questionnaire items, item-level results were averaged to the question level for reporting.

The results presented in this section focus on three key empirical dimensions. First, response quality is examined through item-level missingness patterns, providing an indication of the practical feasibility of data collection in LMIC settings. Second, distributional characteristics are summarized to identify variables with limited variability that may constrain analytical usefulness. Third, associations with student achievement are reported as descriptive indicators

of empirical relevance, based on weighted regression models which were aggregated to the school level in all assessments to assure comparability of the procedures applied and expressed as explained variance (R^2).

Note that these results do not constitute full psychometric validation. Instead, they provide a pragmatic screening of empirical properties to identify variables that may be weak, uninformative, or potentially problematic for cross-country descriptive and policy-oriented analyses. Given that some items were subsequently adapted or reformulated, the reported statistics should be interpreted as indicative of expected behavior rather than definitive item characteristics.

6.6 Empirical Characteristics of Retained Questions

6.6.1 Overview of Empirical Screening Results

Empirical screening results are reported at question level and are based only on the variables corresponding to the questions and response options ultimately selected for further consideration. In most cases, these variables were drawn from the single assessment that provided the selected question. Where a selected question comprised multiple variables, drawn from one or, in rare cases, several source questions, such as the rows of a matrix question or several household possessions, a concern was flagged at the question level if more than 50% of the associated selected variables exceeded the predefined threshold. The percentages reported in Table 6-17 are calculated relative to the number of selected questions included in the empirical screening at each questionnaire level. Detailed item-level results are presented in Table 6-21 (student results), Table 6-22 (teacher results), and Table 6-23 (school results).

Table 6-17: Summary of Empirical Screening Results by Questionnaire Level

Level	N questions	N Questions screened	Missingness concern n (%)	Dominant category concern n (%)	Both concerns n (%)	Low association ($R^2 < 0.01$) n (%)
Student	36	21	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (14.3%)
Teacher	36	22	0 (0.0%)	3 (13.6%)	0 (0.0%)	9 (40.9%)
School	41	31	0 (0.0%)	3 (9.7%)	0 (0.0%)	4 (12.9%)
Total	113	74	0 (0.0%)	6 (8.1%)	0 (0.0%)	16 (21.6%)

Note: Percentages are based on the number of selected questions included in the empirical screening. For selected questions represented by multiple variables, a concern was flagged at the question level if more than 50% of the associated variables exceeded the relevant threshold: missingness >20%, dominant response category $\geq 80\%$, or $R^2 < 0.01$.

The empirical analyses presented in Table 6-17 are restricted to the subset of retained questions with minimal or moderate levels of modification ($MOD \leq 2$). Questions that underwent substantial adaptations were excluded, as changes in wording, structure, or response formats limit the interpretability of empirical characteristics derived from the original items. Even relatively small modifications may influence psychometric properties, and this is particularly relevant for questions classified as having moderate adaptations ($MOD = 2$). In addition, the reported empirical properties are based on a single assessment context and may differ when

the same questions are implemented in other countries or regions. As a result, the findings presented here should be interpreted as providing guidance for selecting and adapting questions for LMIC contexts rather than as definitive evidence for the final questionnaire design. The proportion of questions meeting this criterion varies across questionnaire levels and is highest at the school level, as outlined in the preceding section.

Within this subset, the empirical screening indicates generally high response rates and balanced distributional properties across the majority of retained questions, with only a limited number of items exhibiting potential concerns.

At the student level, none of the questions exceeded the predefined thresholds for missingness or dominant response categories (i.e., $\geq 80\%$ of responses concentrated in a single category). However, three questions (14.3%) showed low associations with student achievement ($R^2 < 0.01$), indicating limited explanatory value for a small subset of constructs.

A similar pattern emerges at the teacher level. While no issues related to missingness were observed, three questions (13.6%) displayed response distributions where $\geq 80\%$ of responses are concentrated in a single category, indicating limited variability. In addition, nine questions (40.9%) exhibited negligible associations with achievement, representing the highest share of low explanatory power across all levels. This comparatively high share is likely influenced by the analytical approach adopted in this study: to ensure comparability across assessments with differing linkage structures, both teacher variables and student outcomes were aggregated to the school level. In the absence of direct teacher–student linkages, this aggregation may reduce the strength of observed relationships, as classroom-level processes are averaged across heterogeneous teaching contexts within schools.

At the school level, response coverage remained high, with no questions affected by missingness concerns. Three questions (9.7%) showed dominant response categories, and four questions (12.9%) demonstrated low associations with achievement. Compared to the teacher level, the proportion of low-association items is notably smaller. Across all levels, no questions were simultaneously affected by both missingness and dominant category issues, indicating that compounded empirical limitations are not a major concern in the current set of retained questions.

Overall, only a small proportion of questions (8.1%) exhibited distributional limitations, while approximately one-fifth (21.6%) showed negligible associations with student achievement.

6.6.2 Association with Achievement

Associations between background variables and achievement are based only on the subset of assessments in which variables were available and are averaged across reading and mathematics domains where applicable. Given these constraints, as well as the aggregation of data to the school level and the reliance on a limited set of assessment contexts, the results should be interpreted as indicative rather than definitive and are not used in the final item suitability assessment.

Associations between background variables and student achievement vary across questionnaire levels. At the student level, 3 out of 21 items (14.3%) with $MOD < 3$ exhibit low associations ($R^2 < 0.01$). A higher proportion is observed at the teacher level, where 9 out of 22 items (40.9%) fall below this threshold. This comparatively high share is likely influenced by the underlying study design, as the availability and structure of student–teacher linkages differ across assessments. To ensure comparability, teacher variables and student outcomes were

aggregated to the school level, which may reduce the strength of observed relationships. At the school level, 4 out of 31 items (12.9%) show low associations with achievement.

6.7 Qualitative Review of Retained Questions

6.7.1 Analytical Scope and Data Sources

Following the empirical screening of retained questions, a structured qualitative review was conducted to incorporate contextual expertise from LMIC representatives. While the empirical analyses provide information on response behavior and statistical properties, the qualitative review focuses on the practical suitability of the proposed questionnaire items, including clarity and their ability to differentiate meaningfully across response categories, as well as their policy relevance of the proposed items within national contexts. The qualitative results therefore provide complementary evidence to support the interpretation of empirical findings and to inform the final selection and adaptation of questionnaire items.

Qualitative evidence was collected through structured ratings and open-ended feedback provided by six NRCs and their teams participating in the LaNA 2023 Link Study. While most participating countries provided feedback across questionnaire levels, the representative of Pakistan contributed only to the teacher questionnaire, and some NRCs did not respond to all questions. As a result, the number of responses per question varies, with typically between four and six countries contributing ratings.

6.7.2 Perceived Relevance and Acceptance of Retained Questions

For each question, NRCs provided categorical ratings across three levels, which were subsequently transformed into a numeric scale ranging from 0 to 2, with 2 representing the most favorable assessment. Average ratings were then computed for each question, and values below 1.5 were flagged as potential concerns. In contrast to the empirical screening, qualitative ratings were evaluated for the full set of retained questions, as expert judgments remain informative even for substantially adapted items. The results of this evaluation are shown in Table 6-18 below.

Table 6-18: Summary of NRC Acceptance Ratings by Questionnaire Level

Level	N questions	Clarity < 1.5 N (%)	Difficulty < 1.5 N (%)	Policy Relevance < 1.5 N (%)
Student	36	2 (5.6%)	8 (22.2%)	2 (5.6%)
Teacher	36	0 (0.0%)	1 (2.8%)	0 (0.0%)
School	41	0 (0.0%)	1 (2.4%)	0 (0.0%)
Total	113	2 (1.8%)	10 (8.8%)	2 (1.8%)

The qualitative ratings indicate a generally high level of acceptance of the retained questions across all questionnaire levels. Issues related to clarity are rare, with only two student-level questions (5.6%) falling below the threshold and no concerns identified at the teacher or school levels. Similarly, perceived policy relevance is consistently high, with only two student-level questions receiving lower ratings, suggesting strong alignment of the retained questions with policy-relevant constructs in LMIC contexts.

Perceived difficulty or feasibility shows a somewhat different pattern. While concerns remain limited overall, they are more pronounced at the student level, where eight questions (22.2%) fall below the threshold. At the teacher and school levels, only isolated cases are observed. This suggests that challenges related to comprehension, response processes, or contextual feasibility are more likely to arise for student-reported items, where there may be challenges in matching wording to what students at the end of primary education can reliably understand and respond to.

The qualitative results suggest that the retained questions are generally well understood, contextually appropriate, and relevant for policy-oriented analyses. The observed differences across questionnaire levels highlight the importance of considering respondent group and construct complexity when adapting and implementing questions in LMIC contexts.

6.7.3 Overall Evaluation of Suggested Questions and Indicators

In addition to item-level ratings, NRCs provided an overall evaluation of the proposed questionnaire item sets with respect to coverage of key constructs, potential missing elements, and national acceptability. For reporting purposes, the three-category ratings (high, medium, low) were converted to a numeric scale (2, 1, 0) and averaged across respondents for each questionnaire level. Overall feedback for this component was received from 5 of the 6 LaNA 2023 Linking Study countries. Table 6-19 summarizes these average ratings.

Table 6-19: Overall Evaluation of Suggested Questions and Indicators (Average Ratings)

Level	Coverage (avg)	Missing elements (avg)	Acceptability (avg)
Student	1.8	1.2	2.0
Teacher	1.8	1.2	2.0
School	1.8	0.8	2.0

The results indicate a high level of acceptance of the proposed questionnaire item sets across all levels, with average ratings of 2.0 for national acceptability, indicating that all responding countries rated this dimension at the maximum level. Coverage of key constructs is consistently rated as high, suggesting that the selected indicators and questions are perceived to capture the most relevant domains for LMIC contexts. The identification of missing elements indicates that some gaps remain, however, particularly at the school level, where lower average ratings suggest that additional constructs or refinements may be needed. For student and teacher questionnaires, concerns regarding missing elements are present but less pronounced.

6.7.4 Missing Elements as Reported by LaNA 2023 Linking Study NRCs

In addition to structured ratings, NRCs were asked to identify potential missing elements in the suggested questions and indicators. Feedback for this component was received from five of the six participating countries and was reviewed across questionnaire levels. Responses were provided in both English and French and were consolidated into thematic areas.

At the student level, suggested additions primarily relate to learning resources and contextual constraints, including individual access to (electronic) textbooks, educational materials, and

ICT, as well as asking more comprehensively about child labor. These elements point to the need of some countries in capturing student-level context more comprehensively.

At the teacher level, feedback emphasizes pedagogical processes and professional conditions, including instructional support from school leadership, formative assessment practices, participation in professional or pedagogical networks, but also teacher's travel time. These suggestions highlight the importance of more fully capturing school-level processes and teaching conditions.

At the school level, missing elements relate mainly to infrastructure, environmental conditions, and governance. These include the need to address more comprehensively the access to water and sanitation, physical school conditions (e.g., fencing, shade), as well as organizational aspects such as multigrade teaching, double-shift systems, and school management structures.

Across all levels, the identified gaps predominantly concern contextual and process-related dimensions, suggesting that these aspects may be somewhat less represented in the current selection of questions and indicators.

These findings should be interpreted as indicative but may provide useful guidance for the continued development and adaptation of questions and indicators in LMIC contexts.

6.7.5 NRC Comments Regarding Individual Questions

In addition to the structured ratings, NRCs provided item-specific comments on wording, response options, and contextual fit. These comments were highly specific to individual questions and countries and are therefore not reported in full here. Instead, they were reviewed for recurring patterns across questionnaire levels.

At the student level, comments of individual reviewers mainly concerned the need for clearer and more age-appropriate wording, especially where students are asked to report on household characteristics, parental background, or concepts that may be difficult to understand reliably at the end of primary education. Further comments emphasized the need for country-specific adaptation of categories and examples, particularly for languages, qualifications, and household resources.

At the teacher level, comments focused primarily on adapting wording and response formats more closely to national and professional contexts. Recurring issues included requests to replace international labels with locally meaningful qualification categories, to clarify ambiguous stems, to simplify overlapping response categories, and to split complex matrix questions into more manageable item formats. Several comments also highlighted the importance of reflecting context-specific teaching arrangements, such as multigrade instruction, contract types, supervision structures, and school or classroom library models.

At the school level, comments similarly emphasized the need for contextual adaptation, especially regarding administrative categories, infrastructure, and organizational arrangements. Recurrent themes included alignment with local qualification systems, school ownership and governance structures, infrastructure and school facilities, textbook provision, double-shift or multigrade arrangements, and the framing of school rules, safety, and repetition policies. As at the teacher level, some NRCs suggested separating longer matrix-style questions into simpler item-by-item formats.

These item-specific comments suggest that the main challenges lie less in the relevance of the selected constructs than in achieving an appropriate balance between maintaining international comparability and ensuring that wording, examples, response formats, and institutional categories are clearly understood within national and respondent-specific contexts.

6.7.6 Summary of the Qualitative Findings

The qualitative findings indicate a high level of policy relevance and acceptance of the suggested questions and indicators across participating countries, with limited concerns regarding clarity and feasibility. The identification of missing elements and the pattern of item-specific comments reflect the importance of considering further refinement of item formulations and contextual specifications for certain items with lower ratings. Individual comments often point to the challenge of balancing international comparability with clear and meaningful wording understood in the individual country context. They also suggest that, in LMIC contexts, internationally comparable instruments may benefit from the inclusion of complementary, context-specific elements, even if these are not uniformly applicable across all LMICs countries.

6.8 Item Suitability Assessment

To integrate the empirical and qualitative evidence presented in the previous sections, an item suitability assessment was conducted at the question level. This assessment combines two complementary sources of information:

- (i) the psychometric screening results derived from existing large-scale assessment data in LMIC contexts, and
- (ii) expert review feedback provided by LMIC representatives.

The objective of this approach is to provide a structured and transparent basis for identifying questions that may require revision prior to instrument inclusion or pilot testing.

Psychometric suitability was determined based on omission rates and response concentration patterns, aggregated at the question level using a majority rule across underlying variables. Expert-based suitability was derived from ratings of clarity and differentiation, with questions flagged where average ratings indicated potential concerns. Items that were substantially modified or newly developed were excluded from the combined assessment due to the lack of comparable empirical evidence.

Table 6-20 presents the classification rules used to determine the overall item suitability status based on the modification level, expert flagging (NRC concern), and psychometric flagging (psychometric concern). Expert and psychometric ratings were applied only to items with modification levels below 3. Items with modification levels below 3 were classified as 'Suitable' when neither expert nor psychometric concerns were identified, 'Consider Revision' when either an expert or psychometric concern was flagged, and 'Strongly Consider Revision' when both concerns were flagged. Items with a major modification (modification level 3) were assigned the status 'Validation', and additionally 'Validation + Consider Revision' when an expert concern was identified.

In addition, information on policy relevance and associations with student achievement is reported to support interpretation and prioritization of items. However, these dimensions were

not included in the suitability classification, as policy relevance reflects broader substantive importance rather than measurement quality, and associations with achievement are subject to the methodological limitations outlined in chapter 6.5.

Table 6-20: Item Suitability Categorization

MOD	NRC (Expert Concern)	PSY (Psych. Concern)	Final Status
< 3	No	No	Suitable
< 3	Yes	No	Consider Revision
< 3	No	Yes	Consider Revision
< 3	Yes	Yes	Strongly Consider Revision
3	No	NA	Validation
3	Yes	NA	Validation + Consider Revision

6.8.1 Overview of Item Suitability Findings

Table 6-21 to Table 6-23 summarize the detailed results of the qualitative and psychometric item review procedures for the student-, teacher-, and school-level questionnaires. In addition to the final suitability classification, the tables include information on policy relevance, clarity, differentiation, expert concerns, psychometric concerns, modification levels, and selected descriptive screening indicators. To facilitate interpretation, the final suitability classifications are color-coded as follows: suitable items are highlighted in green, items requiring validation in yellow, items suggested for revision in orange, and items requiring both validation and revision in red.

Table 6-21: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (Student Level)

ID	Indicator	Stem	Policy Relevance	Clarity	Differentiation	Expert Concern	N Vars	N Issues	N Omissions	N Dominant	Mean R ² (assoc.)	Mod. Level	Psych. Concern	Overall Suitability
1-A	Gender	Which of these describes you?	2	1.5	2		1	0	0	0	0.010	2		Suitable
2-A	Age	How old are you?	2	2	2		2	0	0	0	0.019	1		Suitable
3-A	Language experience	How often do you speak the <language of the test> at home?	2	2	2		1	0	0	0	0.002	1		Suitable
3-B	Language experience	What language do you speak at home most of the time?	2	1.8	1.75		1	0	0	0	0.080	1		Suitable
4-A	Immigration Status	Where were you born?	1.8	2	1.5		1	1	0	1	0.007	3		Validation
5-A	Parents' literacy	How many adults in your home can read or write?	2	1.4	1.5	X	7	4	4	0	0.006	3		Validation and Consider Revision
6-A	Parents' Highest formal education	What is the highest level of education that your father (or male guardian) /mother (or ...	1.8	1.4	1.25	X	2	0	0	0	0.087	2		Consider Revision
7-A	Participation in formal pre-primary education	Did you go to a preschool or kindergarten before starting primary school?	2	2	1.75		1	0	0	0	0.038	2		Suitable
8-A	Grade repetition	Have you ever repeated a grade?	2	2	1.75		1	0	0	0	0.066	1		Suitable
9-A	Absenteeism	About how often are you absent from school?	2	1.75	1.67		1	0	0	0	0.030	2		Suitable
9-B	Absenteeism	About how often do you arrive late to school?	2	1.8	1.75		1	0	0	0	0.005	3		Validation
10-A	Household composition (number of people living at home)	How many people currently live in your home? (Including yourself.)	2	2	2		1	0	0	0	0.007	2		Suitable
10-B	Household composition (number of people living at home)	How many siblings (brothers and sisters) do you have?	1.4	1.8	1.75		1	0	0	0	0.053	3		Validation
11-A	Number of books in the home	How many books are there in your home? (Do not count schoolbooks, newspapers, magazines)	2	1.75	1.67		1	0	0	0	0.067	2		Suitable
12-A	Household possessions and services	Do you have any of these things at your home? a) Running water b) Electricity c) Refrig...	1.75	1.75	1.75		5	1	0	1	0.020	3		Validation
13-A	Type, material of house	What are the walls of your home mostly made of?	1.4	2	1.25	X	2	1	1	0	0.037	3		Validation and Consider Revision
14-A	Parental involvement in learning (incl. help with schoolwork/ Tuition)	How often does someone at home help you with your homework?	1.67	2	1	X	2	0	0	0	0.019	3		Validation and Consider Revision
15-A	Assignment of Homework	About how often do you get homework? a) In mathematics b) In <language of test>/ reading	1.8	1.8	0.75	X	1	0	0	0	0.032	2		Consider Revision
16-A	Enjoyment of reading/ mathematics	How much do you agree with these statements about reading? a) I like talking about what...	1.67	2	2		4	0	0	0	0.007	2		Suitable
16-B	Enjoyment of reading/ mathematics	How much do you agree with these statements about learning mathematics? a) I enjoy lear...	2	2	2		4	0	0	0	0.014	2		Suitable
17-A	Confidence in reading/ mathematics	How much do you agree with these statements about reading? a) I usually do well in read...	1.75	2	2		2	0	0	0	0.030	2		Suitable
17-B	Confidence in reading/ mathematics	How much do you agree with these statements about mathematics? a) I usually do well in ...	1.75	2	2		5	0	0	0	0.059	2		Suitable
18-A	Nutrition, meals	Which meals do you eat on a normal school day? a) Breakfast b) Lunch c) Dinner	2	1.8	1.75		3	3	0	3	0.021	3		Validation
18-B	Nutrition, meals	How often do you have trouble paying attention in school because of these things? a) I ...	2	1.75	1.33	X	1	0	0	0	0.007	3		Validation and Consider Revision
19-A	Physical and mental health (incl. Anxiety, Depression)	How often do you have trouble paying attention in school because of these things? a) I ...	2	1.75	1.33	X	1	0	0	0	0.015	3		Validation and Consider Revision
19-B	Physical and mental health (incl. Anxiety, Depression)	Do you agree with these statements about your health: a) I can clearly see the words th...	2	2	1.75		5	5	0	5	0.032	3		Validation
20-A	Well being at school	How much do you agree with these statements about your school? a) I like being in schoo...	2	2	2		3	1	0	1	0.010	1		Suitable
21-A	Classroom climate	How often do these things happen in your class? a) Students don't listen to what the te...	2	2	2		11	0	0	0	0.019	2		Suitable
22-A	Experience of bullying/ harassment	During this year, how often have other students from your school done any of the follow...	2	2	2		10	3	0	3	0.046	1		Suitable
23-A	Access to books from community/school/class library	How often do you borrow books (including ebooks) from your <class/ school or local libr...	2	2	2		1	0	0	0	0.016	1		Suitable
24-A	House chores/ Work	How often do you do these things outside of school? a) Do house chores (e.g., cleaning,...	2	2	2		6	0	0	0	0.054	2		Suitable
25-A	Commuting to school	How long does it take you to get to school each day?	2	2	2		1	0	0	0	0.030	1		Suitable
26-A	Language of class instruction	How often does your reading/ language teacher use a language different from the <langua...	2	1.8	1.25	X	3	0	0	0	0.009	3		Validation and Consider Revision
26-B	Language of class instruction	Which other language does your language teacher use the most when teaching?	1.75	1.75	1.25	X						3		Validation and Consider Revision
27-A	Perception of Instruction	How much do you agree with these statements about your reading lessons? a) I know what ...	2	2	2		3	1	0	1	0.024	3		Validation
27-B	Perception of Instruction	How much do you agree with these statements about your mathematics lessons? a) I know w...	2	2	2		3	1	0	1	0.021	3		Validation

Table 6-22: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (Teacher Level)

ID	Indicator	Item	Policy Rele- vance	Clarity	Differentiation	Expert Concern	N Vars	N Issues	N Omissions	N Dominant	Mean R ² (assoc.)	Mod. Level	Psych. Concern	Overall Suitability
1-A	Teacher Gender	Which of these describes you?	2	1.75	1.75		1	1	0	1	0.022	2	X	Consider Revision
2-A	Teacher Age	How old are you?	2	2	2		1	0	0	0	0.013	1		Suitable
3-A	Years of Teaching	By the end of this school year, how many years will you have been teaching altogether?	2	1.8	1.8		1	0	0	0	0.011	1		Suitable
3-B	Years of Teaching	By the end of this school year, how many years will you have been teaching at this school?	2	1.75	1.75		1	0	0	0	0.014	3		Validation
4-A	Highest level of education	What is the highest level of formal education you have completed?	2	2	2		1	0	0	0	0.030	1		Suitable
5-A	Professional Teaching Qualification	Do you have a formal teaching qualification that qualifies you to teach in schools? (e....	2	2	1.83		1	1	0	1	0.014	2	X	Consider Revision
6-A	Subject-Matter Assignments (general /in sel. class)	Which of the following two subjects do you teach in the selected class? a) Mathematics ...	2	2	1.8		3	3	0	3	0.009	2	X	Consider Revision
7-A	Grades/ Classes being taught	Which <grade(s) or class level(s)> do you currently teach in this school?	2	2	1.83		1	1	0	1	0.010	3		Validation
8-A	Initial Teacher Education (duration/quality/Experiences)	Did you complete any pre-service teacher training? (This may have been completed as par...	2	1.83	1.67		1	0	0	0	0.002	3		Validation
9-A	Area(s) of Post-secondary Study / Teacher	During your <post-secondary> education, what was your major or main area(s) of study? (...)	2	1.8	1.6		6	1	0	1	0.006	2		Suitable
9-B	Area(s) of Post-secondary Study / Teacher	If your major or main area of study was education, did you have a specialization in any...	2	2	1.8		4	0	0	0	0.005	1		Suitable
10-A	Professional development (duration, quality)	In the past two years, have you participated in formal professional development in the ...	2	1.83	1.5		4	0	0	0	0.013	3		Validation
10-B	Professional development (duration, quality)	If you answered yes to any of the activities above, to what extent did these formal pro...	2	2	1.4	X	1	0	0	0	0.002	2		Consider Revision
11-A	Confidence in teaching (methods)	How confident are you in doing each of these things when teaching this class? a) Having...	2	2	1.83		5	0	0	0	0.014	2		Suitable
12-A	Job satisfaction (incl. General working conditions)	How often do you have these feelings about being a teacher? a) I am content with my pro...	2	2	1.8		7	0	0	0	0.009	1		Suitable
13-A	Teacher Absence	During the last month, did any of the following situations prevent you from going to wo...	2	1.83	1.67		15	11	0	11	0.003	3		Validation
14-A	Type of contract with school/ Position	How long is your employment contract with this school?	2	2	2		1	0	0	0	0.028	3		Validation
15-A	Additional paid jobs	In addition to your regular teaching at this school, how many hours per week do you spe...	2	2	1.83		2	0	0	0	0.007	2		Suitable
16-A	Instructional Resources and Classroom Infrastructure Inventory	Which of the following items do you have in your classroom? a) Teacher's guide b) Elect...	2	1.83	1.67		2	0	0	0	0.017	3		Validation
16-B	Instructional Resources and Classroom Infrastructure Inventory	How does your school make textbooks available to students in this class? a) Mathematics...	2	2	2		1	0	0	0	0.046	3		Validation
17-A	School- or Classroom library and use	How many books with different titles are in your classroom library or reading corner?	2	2	1.83		2	0	0	0	0.016	3		Validation
17-B	School- or Classroom library and use	Can the students borrow books from the classroom library or reading corner to take home?	2	2	2		1	0	0	0	0.018	1		Suitable
18-A	EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)	How would you characterize each of the following within your school? a) Teachers' under...	2	1.83	1.67		4	0	0	0	0.019	1		Suitable
19-A	EAS_Parental Support	How would you characterize each of the following within your school? a) Parental involv...	2	2	1.83		4	0	0	0	0.039	1		Suitable
20-A	SOS_Personal Safety (including bullying)	How often have these things happened at your school in the past year? a) A student thre...	2	2	1.83		1	1	0	1	0.015	3		Validation
20-B	SOS_Personal Safety (including bullying)	How much do you agree or disagree with the following statements about your current scho...	2	2	2		4	0	0	0	0.013	2		Suitable
21-A	SOS_Student Behavior & Respect	How much do you agree or disagree with the following statements about your current scho...	2	2	1.83		3	0	0	0	0.015	3		Validation
21-B	SOS_Student Behavior & Respect	How often do the following situations occur in the <classroom assessed>? a) When the cl...	2	2	2		3	0	0	0	0.009	2		Suitable
22-A	External Supervision & Inspection	Since the beginning of this school year, how often have the following people visited yo...	2	2	1.83		3	0	0	0	0.007	3		Validation
23-A	Class size/ number of students on day of test	How many students are in this class? a) How many students are in this class in total? b...	2	2	1.83		2	0	0	0	0.045	3		Validation
24-A	Teaching conditions/ Teaching Constraints	In your view, to what extent do the following limit how you teach this class? a) Studen...	2	2	1.83		9	0	0	0	0.013	1		Suitable
24-B	Teaching conditions/ Teaching Constraints	To what extent do the following limit how you teach this class? a) Lack of basic resour...	2	2	2							3		Validation
25-A	Language of Instruction	What languages do you use when teaching your students in this class?	2	1.83	1.83		1	0	0	0	0.048	1		Suitable
26-A	Instructional time (Language/ Reading or mathematics)	In a typical week, how much time do you spend teaching mathematics to the students in L...	2	1.83	1.83		1	0	0	0	0.014	1		Suitable
26-B	Instructional time (Language/ Reading or mathematics)	In a typical week, how much time do you spend on <language of test> language instructio...	2	1.8	1.8		1	0	0	0	0.008	1		Suitable
27-A	Instructional activities (general, Language/ Reading, Math)	How often do you do the following in teaching this class? a) Relate the lesson to stude...	2	2	2		7	2	0	2	0.008	2		Suitable

Table 6-23: Detailed Psychometric Screening and Qualitative Findings on Questionnaire Level (School Level)

ID	Indicator	Item	Policy Rele- vance	Clarity	Differentiation	Expert Concern	N Vars	N Issues	N Omissions	N Dominant	Mean R ² (assoc.)	Mod. Level	Psych. Concern	Overall Suitability
1-A	Gender	Which of these describes you?	2	2	1.67		1	0	0	0	0.014	2		Suitable
2-A	Age	How old are you?	2	1.75	1.67		1	0	0	0	0.064	1		Suitable
3-A	Years of service, experience in education	By the end of this school year, how many years will you have been a principal altogether?	2	2	1.67		1	0	0	0	0.019	1		Suitable
3-B	Years of service, experience in education	By the end of this school year, how many years will you have been a principal at this s...	2	2	1.67		1	0	0	0	0.016	1		Suitable
4-A	Highest level of education	What is the highest level of formal education you have completed?	2	2	2		1	0	0	0	0.012	2		Suitable
5-A	Professional development	Have you completed any formal professional training in education administration or mana...	2	2	2		3	0	0	0	0.008	3		Validation
6-A	School ownership (private/public/religious)	Which of these options best describes your school?	2	2	1.75		1	1	0	1	0.090	3		Validation
6-B	School ownership (private/public/religious)	What kind of organization manages your school?	2	1.8	1.5		1	0	0	0	0.136	3		Validation
7-A	School type/structure (primary/ secondary/language)	Which of the following levels of education does your school provide?	2	1.8	1.75		4	1	0	1	0.121	1		Suitable
8-A	School location (urban/rural)	How many people live in the city, town, or area where your school is located?	1.8	1.8	1.75		1	0	0	0	0.062	1		Suitable
8-B	School location (urban/rural)	Which best describes the immediate area in which your school is located?	2	2	2		1	0	0	0	0.073	1		Suitable
9-A	SES Backgrounds of students	Approximately what percentage of students in your school have the following backgrounds...	2	1.8	1.5		2	0	0	0	0.092	2		Suitable
10-A	School size (total enrollment of boys and girls and at target grade)	What is the total enrollment of students in your school as of <first day of month test...	2	2	1.75		3	0	0	0	0.091	3		Validation
11-A	Enrollment & number of classes in target grade	What is the total enrollment of <target grade> students in your school as of <first day...	2	2	2		1	0	0	0	0.037	3		Validation
12-A	Language of test/ language of instruction	Approximately what percentage of students in your school have <language of test> as the...	2	2	2		1	0	0	0	0.057	1		Suitable
13-A	Number of teachers/ Staff	How many teachers are at your school? (A full-time teacher is employed at least 90% of ...	2	2	1.75		1	0	0	0	0.041	3		Validation
14-A	School Educational Resources except books (and affection by shortage)	How much is your school's capacity to provide instruction affected by a shortage or ina...	2	2	2		1	0	0	0	0.012	2		Suitable
15-A	School human resources (and affected by shortage)	How much is your school's capacity to provide instruction affected by a shortage or ina...	2	2	2		1	0	0	0	0.026	2		Suitable
16-A	School Library (equipment and use)	Approximately, how many books does your school library have?	2	2	2		1	1	0	1	0.032	3		Validation
17-A	Internet	Does your school have the following resources? a) Reliable internet connection availabl...	2	2	2		2	0	0	0	0.040	2		Suitable
18-A	School facilities (incl. Toilets SDG)	Does your school have the following resources? a) Electricity b) Safe drinking water av...	2	2	2		9	3	0	3	0.074	3		Validation
19-A	Classroom facilities	Does your school have the following classroom resources? a) Sufficient chairs or benche...	2	1.8	1.75		3	2	0	2	0.012	3		Validation
20-A	Availability of textbooks	How does your school make textbooks available to students in <target grade>? a) Mathema...	2	2	2		2	2	0	2	0.026	2	X	Consider Revision
20-B	Availability of textbooks	How much is your school's capacity to provide instruction affected by a shortage or ina...	2	2	2		1	0	0	0	0.007	2		Suitable
21-A	Provision of additional support in reading or math	Does your school provide extra classes or tutoring for students who need additional sup...	2	2	2		1	0	0	0	0.011	2		Suitable
22-A	Digital learning resources	How many digital devices does your school have for use by <target grade> students?	1.75	1.5	1.33	X	1	0	0	0	0.027	3		Validation and Consider Revision
23-A	School Year Structure (beginning/end of school year, operational days)	For the <target grade> students in your school: How many days per year is your school s...	2	1.67	1.5		1	0	0	0	0.009	1		Suitable
23-B	School Year Structure (beginning/end of school year, operational days)	In one calendar week, how many days is the school open for instruction?	2	2	2		2	0	0	0	0.027	1		Suitable
24-A	School Year Structure (instructional time per day)	What is the total instructional time, excluding breaks, on a typical day?	2	2	2		1	0	0	0	0.004	1		Suitable
25-A	Shifts	Does your school have double shifts or double sessions? (one group of students attends ...	2	1.6	1.75		1	1	0	1	0.005	1	X	Consider Revision
26-A	Days school closed (Covid, strikes, etc.)	Within the last 12 month, about how many days was your school closed due to unforeseen ...	2	2	2		1	0	0	0	0.040	1		Suitable
27-A	Teacher job satisfaction	How would you characterize each of the following within your school? a) Teachers' job s...	2	2	2		1	0	0	0	0.028	1		Suitable
28-A	EAS_ Instructional and Curricular Leadership and Staff Cohesion	How would you characterize each of the following within your school? a) Teacher's under...	2	1.75	1.67		3	0	0	0	0.023	1		Suitable
29-A	EAS_ Parental Support	How would you characterize each of the following within your school? a) Parental suppor...	2	2	2		2	0	0	0	0.038	1		Suitable
30-A	EAS_ Student Engagement and Peer Culture	How would you characterize each of the following within your school? a) Students' regar...	2	2	2		2	0	0	0	0.035	1		Suitable
31-A	SOS_ Personal Safety / BULLYING (incl. Teachers)	To what degree is each of the following a problem among <target grade> students in your...	2	2	2		3	0	0	0	0.027	2		Suitable
32-A	SOS_ School rules and enforcement	To what degree is each of the following a problem among <target grade> students in your...	2	2	2		2	0	0	0	0.011	1		Suitable
33-A	SOS_ Student Behavior & Respect	To what degree is each of the following a problem among <target grade> students in your...	2	2	2		5	0	0	0	0.015	1		Suitable
34-A	Policies on grade retention/ academic support & digital devices	What are your school's policies regarding <grade repetition>? a) Students may <repeat a...	1.8	2	2		6	2	0	2	0.019	1		Suitable
35-A	Teacher Absenteeism, Teacher availability, Issues at school	To what degree is each of the following a problem among teachers in your school? a) Arr...	2	2	2		2	0	0	0	0.025	2		Suitable
36-A	Nutrition, meals provided by school	Does your school provide free meals for students? a) Breakfast b) Lunch	2	2	2		2	2	0	2	0.019	1	X	Consider Revision

6.8.2 Item Suitability Patterns

Item suitability results are reported at the question level and are based on a combined evaluation of expert judgment (NRC), psychometric screening results (PSY), and the degree of modification (MOD). For items with minimal or moderate adaptations (MOD < 3), suitability classifications reflect the combined presence of expert and psychometric concerns. For substantially redeveloped items (MOD = 3), original psychometric evidence was considered no longer transferable; these items were therefore evaluated based on expert judgment only and classified as requiring empirical validation.

Detailed item-level results are presented in the consolidated overviews presented in Table 6-21 (student results),

Table 6-22 (teacher results), and Table 6-23 (school results), while Table 6-24 summarizes the suitability patterns across questionnaire levels.

Table 6-24: Distribution of Item Suitability by Questionnaire Level

Level	Items (total)	Suitable N (%)	Consider Revision N (%)	Strongly Consider Revision N (%)	Requires Validation N (%)	Requires Validation and Revision N (%)
Students	36	19 (52.8%)	2 (5.6%)	0 (0.0%)	8 (22.2%)	7 (19.4%)
Teachers	36	18 (50.0%)	4 (11.1%)	0 (0.0%)	14 (38.9%)	0 (0.0%)
Schools	41	28 (68.3%)	3 (7.3%)	0 (0.0%)	9 (22.0%)	1 (2.4%)
Total	113	65	9		31	8

6.8.2.1 Interpretation and Patterns

The suitability analysis distinguishes between items for which no concerns were raised, those that may require revision - either due to concerns raised by experts or psychometric concerns should be considered, and those that require empirical validation due to substantial modifications. Across all questionnaire levels, a majority of items are classified as suitable (i.e., suggested for further use without modification), indicating that a substantial share of the adapted questions can be used without further changes.

At the student level, just over half of the items (52.8%) are considered suitable, while a relatively large proportion (41.7%) fall into categories requiring validation. Within this group, nearly half (19.4% of all student items) are additionally flagged by experts, indicating a possible need for revision and targeted piloting. The large proportion of items requiring validation indicates that many of the originally selected questions underwent comparatively extensive adaptations to meet the criteria for inclusion in an internationally comparable assessment at the end of primary education. General terms or simplifications used by expert reviewers may, in some cases, have raised concerns among country experts regarding the extent to which questions remain clearly understandable in their contexts.

A similar pattern emerges at the teacher level. While the proportion of suitable items (50.0%) is comparable to the student questionnaire, the share of items with substantial modification and therefore requiring validation is higher (38.9%). However, only 4 of the originally retained items (11.1%) were flagged by country experts for possible revision, and none of these were among the substantially modified items.

In contrast, the school questionnaire shows the highest level of overall stability. More than two-thirds of the items (68.3%) are classified as suitable, with about a quarter of the items requiring validation (23.0%), and only about 10% of items recommended for revision (three less modified items and one substantially modified item). This indicates that the school-level constructs are comparatively well aligned across frameworks and required fewer substantial adaptations.

Across all levels, the share of items for which revision should be considered—combining “Consider Revision” and “Requires Validation and Revision”—is modest overall (17 of 113; 15.0%), with higher shares at the student level and lower at the school level. In parallel, a substantial proportion of items are substantially modified (MOD = 3) and therefore require empirical validation regardless of revision needs (overall 39 of 113; 34.5%), with the highest concentration in the teacher questionnaire (14 of 36; 38.9%).

Among the retained items (MOD<3) classified as “Consider Revision”, concerns were predominantly driven by 3 cases flagged by expert judgment and 6 cases flagged based on psychometric screening results, with no overlap between both sources of concern.

6.8.2.2 Interpretation Notes

The classification presented in this section should be interpreted as indicative rather than definitive. In particular, items classified as requiring validation (MOD = 3) reflect a lack of transferable psychometric evidence rather than confirmed measurement issues and should be prioritized for empirical testing in future administrations. Similarly, expert judgments are based on the adapted item formulations and represent anticipated suitability across contexts. In addition, the expert review was conducted with a limited number of countries that participated in the LaNA 2023 Linking Study (five for the teacher questionnaire and six for the student and school questionnaires), which may somewhat limit the extent to which the identified concerns generalize across a broader set of contexts.

6.8.2.3 Overall Summary

Overall, the suitability analysis indicates that a majority of items can be suggested for further use without modification, while a substantial share (particularly at the student and teacher levels) require empirical validation due to substantial adaptations. The relatively small number of items flagged for revision (maybe except for the student level where about half of the substantially adapted items are flagged for considering revision) indicates that the adaptation process was mostly effective in preserving content properties and in meeting regional needs. The presence of a sizeable group of redeveloped items reinforces that follow-up validation efforts are needed to ensure the robustness and comparability of final instruments created from the suggested set of items.

7 Discussion and Implications

7.1 Synthesis of Findings and Answer to Research Questions

This section addresses the research questions (see section 3.2), summarizes the findings of the study, and shows how the analysis progressed from framework comparison to indicator identification and, finally, item evaluation. This approach follows the analytical sequence used in the study: i.e., frameworks → indicators → operationalization → evaluation, to illustrate how a robust and policy-relevant set of context indicators for LMICs can be constructed.

7.1.1 Common and Distinct Themes Across Assessment Frameworks (RQ1)

Comparative analysis of the assessment frameworks identified both common elements and systematic differences in the way that learning and its determinants are conceptualized across system levels. These differences were classified according to the analytical project framework (see section 4.2), which differentiated between curricular antecedents, curricular contexts, and curricular content at the school, class, and student levels, as well as between international and regional assessments.

Across almost all frameworks, a common conceptual foundation was evident. Reading and mathematics were included in all frameworks, and each adopted a multi-level perspective, incorporating system, school, classroom, and student-level factors to explain variation in learning outcomes.

Subject coverage was not uniform, however. Science was covered in TIMSS, ERCE, and PISA. SEA-PLM and ERCE also included writing, SEA-PLM included Global Citizenship, and SACMEQ assessed health knowledge.

Learning outcomes were primarily operationalized through student achievement, often complemented by attitudinal constructs, e.g., motivation, engagement, or perceptions of learning. PISA-D identified four “prosperity outcomes” which, in addition to learning outcomes, included educational attainment, engagement at school, and health and well-being.

All frameworks covered students’ experiences at school and student background characteristics, particularly socioeconomic status, typically captured through indicators related to the home environment, for example, parental education and home possessions. School and classroom conditions, including teaching practices and learning environments, were also represented across all frameworks, although the level of detail varied.

There were differences in the way these shared elements were structured. IEA study frameworks from TIMSS and PIRLS provided clear representations of curricular content across the intended, implemented, and attained dimensions. At the system level, these frameworks described curriculum structures and education system features. At the school and classroom levels, the implemented curriculum was operationalized using clearly defined and IRT-validated constructs related to instructional practices (e.g. clarity, engagement), school climate (e.g., school safety, emphasis on academic success, bullying), and learning environments (e.g., home resources for learning, instruction affected by shortages). Student-level indicators focused on attained outcomes, including both achievement and subject-specific attitudes.

Regional assessments SACMEQ, PASEC, ERCE, and SEA-PLM placed less explicit emphasis on curricular content, particularly at the intended and implemented levels. Curricular dimensions were often only implicitly represented or had to be reconstructed from instruments and reporting frameworks. References to teaching and learning processes were often broader and less systematically linked to a curricular model: this resulted in a less explicit representation of the implemented curriculum.

These frameworks extended more strongly toward curricular antecedents (background characteristics, prior conditions, and enabling factors at the system, school, classroom, and student levels that precede educational processes) and contextual conditions (the social environments, organizational arrangements, practices, and behaviors through which teaching and learning take place).

This was reflected in their systematic attention to whether basic conditions for learning were in place, including school infrastructure and service provision, as well as classroom-level constraints such as teacher absenteeism, student truancy, and resource shortages. At the student level, background measures were broadened to capture more directly the living conditions of students, including housing quality, food security, and health-related factors. This emphasis on basic learning conditions reflects the regional assessments' focus on LMICs. Conditions that are often taken for granted in high-income systems become explicit objects of measurement.

Among assessments tailored to LMIC contexts, PISA-D differed from purely regional assessments in its aim to enable comparability across diverse settings. It incorporated context-sensitive measures such as expanded background indicators, well-being, and the inclusion of out-of-school youth, while retaining an explicit conceptual framework in which constructs and their relationships were clearly defined.

The differences reflected underlying assumptions about educational systems and their conditions. IEA frameworks implicitly assumed a baseline level of infrastructure and resource provision, focusing on variation in instructional processes and learning outcomes, while regional frameworks explicitly examined whether such baseline conditions were met. This reflects the heterogeneity and resource constraints characteristic of many LMIC contexts. PISA-D, in turn, acknowledged these constraints while attempting to retain a degree of standardization to support cross-national comparison.

Teachers were recognized as central to classroom processes in all the examined frameworks. Teacher-related constructs were not consistently represented as a distinct analytical level, however. Several studies included dedicated teacher questionnaires covering experience, education, and instructional practices, but these variables were often embedded conceptually within school or classroom domains.

These differences in themes represent a continuum in assessment design, ranging from highly standardized international models to context-embedded regional approaches, with PISA-D (a hybrid framework) combining elements of both.

7.1.2 Identification of Robust and Policy-Relevant Indicators (RQ2)

The identification of indicators followed a structured scoring approach that considered representation across regional assessment initiatives, alignment with SDG priorities, partner

requests, and availability within existing IEA ILSAs. First, questionnaire content from all included assessments was systematically sorted into domains (e.g., student home environment or student preparation for learning) and aligned across instruments. This provided the basis for the construction of indicators (e.g., household composition or language of instruction). More than one indicator could be assigned to one domain. Indicators were then selected based on their recurrence and relevance across LMIC contexts.

Second, criteria relating to policy relevance and psychometric properties were considered during the item-level review. During this phase, specific questionnaire items were selected and (in some cases) adapted to operationalize the selected indicators. These were then evaluated individually using expert judgment from the LaNA 2023 Linking Study NRCs and study contributors, together with available psychometric evidence.

Indicators that appeared consistently across multiple assessment frameworks were considered part of a core set. Their repeated inclusion suggests broad relevance across LMIC contexts and reflects existing empirical evidence linking these constructs to learning outcomes.

Indicators with more limited representation were reviewed individually by project experts, allowing context-specific but policy-relevant indicators to be retained through an “expert override” mechanism.

Out of 184 candidate indicators assessed across student, teacher, and school levels, 90 were retained based on the combined scoring criteria. Approximately one-third of the retained indicators were not part of existing IEA frameworks and rather originated primarily from regional and LMIC-focused assessments.

Across levels, the additional indicators retained through this process were concentrated in domains that corresponded closely to the gaps identified in the framework comparison. For example, indicators related to household conditions, grade repetition, and language of instruction at the student level; workforce conditions, preparation pathways, contract types, and broader employment realities at the teacher level; and organizational features such as shift systems, retention policies, and supplementary support structures at the school level.

This distribution aligns closely with findings from the literature on educational effectiveness in LMIC contexts described in section 4.1.3.3, which details the importance of material living conditions, health constraints, participation patterns, instructional capacity, and school organization as key drivers of learning outcomes.

The results point to structural limitations in the design of global comparative frameworks such as those for TIMSS and PIRLS. These frameworks prioritize cross-national comparability and long-term trend measurement (Mullis, I. V. S., Martin, M. O., & von Davier, M. [Eds.], 2021), leading to the use of standardized proxy indicators that perform well in HIC contexts but may capture limited variance in LMIC settings. For example, commonly used socioeconomic background indices based on home educational resources may explain substantially less variance in achievement in LMIC contexts than in HICs, where such resources are more widely distributed, a limitation noted by Dolata (2008) and Rutkowski & Rutkowski (2013). These studies further argue for more context-sensitive measures and show that adapting socioeconomic indicators to include basic material conditions, such as housing quality or access to essential services, can improve their explanatory power in some LMIC contexts. Likewise, concerns have been raised regarding the cultural comparability and relevance of

standardized socioeconomic background measures developed for HICs, suggesting that the inclusion of regionally adapted indicators may enhance both validity and interpretability.

These findings support a two-tiered approach to indicator selection for comparative assessments in LMIC contexts. A set of core indicators, supported by broad empirical evidence and cross-assessment consistency, provides a stable foundation for international comparability. Additional context-sensitive indicators capture the structural and environmental conditions that influence learning but are less well represented in existing international frameworks. The retained indicator set in this study reflects this dual approach, combining established internationally comparable constructs with targeted extensions that address systematic gaps in existing frameworks. Given the generally lower reading proficiency of students in many LMIC contexts, this dual approach should be accompanied by an accessible questionnaire design, including concise wording, simple response formats, limited cognitive burden, and, where appropriate, administrator support.

7.1.3 Operationalization of Indicators with Existing Questionnaire Items (RQ3)

The operationalization of the selected indicators into questionnaire items followed a “build-on-existing” approach. The goal was to maximize the use of established instruments while ensuring contextual relevance and feasibility across LMIC settings.

In total, the 90 retained indicators were operationalized into 113 questionnaire items, comprising 36 student, 36 teacher, and 41 school-level items. In most cases, a one-to-one correspondence between indicators and questionnaire items was achieved. Where necessary, additional items were included to capture more complex constructs or to differentiate between subject-specific dimensions, particularly for reading and mathematics outcomes.

7.1.3.1 *Origin of Items in the Final Collection*

Whereas the indicators themselves were selected on the basis of their prevalence across assessment instruments, statistical association with achievement, and expert review (see Section 7.1.2), the composition of the final item set—that is, the choice of items selected to operationalize each selected indicator—reflects a deliberate prioritization of established IEA questionnaire content. As a result, most items were derived from IEA assessments, particularly at the school and teacher levels, where constructs were found in this study to be relatively stable and transferable across contexts.

Non-IEA assessments were primarily used to operationalize indicators not sufficiently covered in TIMSS and PIRLS, or LaNA. Indicators capturing LMIC-specific household and schooling conditions were therefore complemented with items from regional instruments.

The resulting, final item set integrates IEA questionnaire content, which provides psychometric validity and cross-national comparability, with items from regional and LMIC-focused assessments, selected for their established validity in LMIC educational contexts.

7.1.3.2 *Cross-Level Differences in Operationalization*

A comparison across levels (student, school, and teacher) highlighted important differences in both the sources used and the extent to which items needed adapting.

At the school level, most indicators could be operationalized using existing IEA items. School-level constructs were found to be more consistently represented across the different assessment frameworks.

At the teacher level, a similar, though more balanced pattern was observed. While many indicators were covered by IEA-based items, additional input from regional assessments was required to capture context-specific aspects of teacher working conditions, employment structures, and preparation pathways.

The student level showed a substantially higher reliance on non-IEA sources and adapted items. This reflected the importance of capturing context-specific aspects of students' realities in LMICs, particularly household conditions, language use, and participation patterns. Several of these constructs, such as parental background and aspects of the home learning environment, are typically collected through parent questionnaires in TIMSS and PIRLS. Where parent questionnaires are not administered, these constructs must rather be collected through student reports, often drawing on items from non-IEA assessments.

7.1.3.3 Adaptation Requirements and Design Constraints

These differences in source reliance were closely mirrored in the adaptation patterns. While school-level items often required only minimal adjustments from the original versions, and teacher-level adaptations were primarily driven by the need to adjust terminology, student-level items required the most extensive modifications. These adaptations were mostly intended to ensure age-appropriate comprehension and contextual validity, including the simplification of language, the adaptation of household and resource categories, and the adjustment of terminology for primary-age respondents. Region-specific wording sometimes required generalization to ensure broader understanding across LMIC contexts, where the same underlying item may be referred to differently in different assessments (e.g., stove, heater, cooker, or fireplace).

The adaptation classification distinguished between minimal, moderate, and substantial modifications. Importantly, higher levels of adaptation did not necessarily indicate a conceptual departure from the original construct. In many cases, substantial modifications were required to translate items from non-IEA sources into a format compatible with IEA questionnaire conventions, including converting response scales, restructuring item formats, or simplifying complex item lists.

Such adaptations may, however, affect item measurement properties. While efforts were made to standardize item formats and ensure consistency, the transfer of items across contexts and formats may affect their psychometric properties. The validity of psychometric properties derived from source assessments may therefore be limited when applied to different LMIC contexts.

7.1.3.4 Implications of Source Diversity for Measurement

The operationalization process was further shaped by differences in questionnaire design across assessments. Regional assessments PASEC and SACMEQ, for example, employ longer and more complex student questionnaires, in some cases including filter questions and skip logic. While these designs may allow for more detailed contextual measurement, they may also increase respondent burden and contribute to higher omission rates, with implications for data quality and comparability.

In addition, the structure of items differed between regional assessments reflecting different approaches to measuring socioeconomic background. In SACMEQ and PASEC, measures focused more frequently on material living conditions and household wealth at the lower end of the distribution, including housing materials and access to essential services. This is consistent

with the composition of these studies, which included a larger share of low-income countries (see Figure 2-1), and therefore required measures sensitive to variation in basic living conditions. In contrast, ILSAs and ERCE placed greater emphasis on home educational resources and other household possessions. Such measures capture specific dimensions of socioeconomic background rather than socioeconomic status in the strict sense, which is commonly operationalized through family income, parental education, and occupational status (Bradley & Corwyn, 2002). The relevance and discriminatory value of particular possession indicators may also vary across countries. These patterns are consistent with the broader literature on socioeconomic stratification, which distinguishes material deprivation from broader socioeconomic resources (e.g., Bourdieu, 1986; Bradley & Corwyn, 2002).

These findings highlight that operationalization was shaped by the conceptual design of each assessment framework and the educational contexts in which it was applied.

Drawing primarily on the globally comparable perspective of IEA assessments, the present project sought to identify a pragmatic compromise across diverse LMIC regions. This involved balancing the need to capture context-specific variation, including conditions at the lower end of the socioeconomic distribution, with the requirement for a shared set of items that remained broadly applicable and comparable across countries.

7.1.4 Item Evaluation and Suitability Review (RQ4)

The final stage of the analytical pipeline combined psychometric screening and expert review to assess the suitability of the operationalized items.

In summary, most items (65/113) were deemed suitable for further use without adaptation. Eleven items, mainly at the student level and often those that needed substantially to be modified (Modification level = 3), were flagged by expert review for potential revision, while six items were flagged based on data-quality concerns (item non-response and low response variance). As the expert- and data-quality flags did not overlap, 17 items were identified for closer review before implementation. An additional 39 substantially adapted items were identified as priorities for empirical validation. These flags are intentionally conservative and do not imply that items are flawed. Rather, they identify items that may benefit from refinement, context-specific adjustment, or further empirical evidence.

A key outcome of this process is an item suitability classification that provides practical guidance for researchers and practitioners implementing the proposed questionnaire item sets or subsets. For each item, the classification reports expert assessments of clarity and differentiation, perceived policy relevance in participating LaNA countries, and empirical indicators such as response distributions and associations with achievement. This multidimensional perspective supports informed decisions about item selection in relation to specific analytical or policy objectives.

These metrics should be interpreted with appropriate caution. Associations with achievement are indicative rather than causal and are subject to limitations related to aggregation, data linkage, and cross-country comparability. Likewise, items classified as requiring validation may reflect the absence of transferable evidence following adaptation rather than confirmed deficiencies in measurement quality.

Overall, the integrated evaluation indicates that a substantial share of items can be used directly or with limited revision, while identifying a clearly defined subset requiring further validation.

7.1.5 Overall Synthesis

Across framework comparison, indicator selection, operationalization, and item evaluation the findings demonstrate that it is feasible to construct a robust and policy-relevant set of context indicators for end-of-primary assessments in LMIC contexts.

Developing such an indicator set is not a purely technical exercise but requires flexibility in how indicators are operationalized. Items often need to be generalized or reformulated to remain understandable and relevant across diverse LMIC contexts. For example, context-specific terms may need to be replaced by broader categories (e.g., combining items such as stove, heater, or fireplace into more general formulations). While this broadening of item wording increases applicability across contexts, it may also reduce precision. This process of ‘widening the context’ is especially evident when adapting items from regional assessments, which often operationalize constructs in ways that reflect geographic or cultural contexts. Expert review and empirical screening are therefore essential in identifying potential measurement issues and ensuring that adapted items remain relevant across contexts.

The findings suggest that learning assessments in LMIC settings should combine a core set of internationally comparable indicators with context-specific indicators. IEA frameworks and instruments provide a strong comparative foundation, particularly for core constructs related to curricular content and established contextual factors; regional and LMIC-focused assessments contribute additional domains and measurement approaches that strengthen the relevance of context questionnaires. These contributions are particularly evident for contextual factors that influence learning, including household environment, language of instruction, participation patterns, and broader measures of health and nutrition, especially at the student level.

7.2 Methodological Reflections and Limitations

The study provides a systematic, albeit exploratory, approach to identifying and operationalizing context indicators for LMIC settings. Several limitations should be considered, however, when interpreting the findings.

The first limitation relates to the availability and transferability of psychometric evidence. The empirical screening of items was based on data from existing regional assessment databases, which are not necessarily representative of all LMIC contexts. Psychometric properties were derived from the original question formulations (including, in some cases, response options that were combined, such as for household possession items) and typically originated from a single regional assessment, which may limit their generalizability across regions. Given that a substantial share of items was adapted during the operationalization process, these properties may not be directly transferable to the revised items or applicable across other regional settings. To address this, items were classified into modification levels, and psychometric properties were reported only for items judged to be sufficiently close to their original formulations. In addition, the analytical approach to empirical validation faced constraints related to data aggregation at the school level and to differences in assessment design across the studies (e.g., availability of teacher-level data, student–teacher linkage designs), which may affect observed relationships and limit comparisons across assessments. The psychometric findings should therefore be interpreted as indicative and require confirmation through further empirical validation. Omission rates may reflect not only the characteristics of individual items but also the overall length and complexity of the questionnaire. The PASEC questionnaire, for example, was relatively long and included several filter questions, which may partly explain high omission rates in certain sections.

The second limitation concerns the scope of expert input. The qualitative review was conducted with representatives from the six LaNA 2023 Linking Study countries (Burkina Faso, Nigeria, Senegal, Pakistan, Egypt, and Palestine), with Pakistan contributing only to the teacher questionnaire review. While this provided valuable insights across Sub-Saharan African and Western Asian contexts, regions such as Latin America and Southeast Asia were not represented.

The third limitation relates to the underlying assessment instruments that formed the basis of this study. The analysis relied on period-specific instruments, primarily reflecting the state of international and regional assessments around 2019 (with the exception of SACMEQ, for which SACMEQ V information was not available; the analysis therefore relied on SACMEQ IV [2013–2014] instruments). While the LaNA 2023 Linking Study instruments were included in the analysis, more recent developments in international assessments, such as ongoing questionnaire revisions for TIMSS 2027 and PIRLS 2026, were considered only indirectly through expert input. Consequently, the analysis does not fully reflect the most recent changes in international and regional assessments. Emerging domains that have become increasingly relevant in LMIC contexts, such as digital learning environments and the role of artificial intelligence in education, were likewise not systematically covered.

A further limitation relates to questionnaire design and implementation. The study focused on item content and structural characteristics, but did not systematically address aspects of visual design and administration procedures, such as the use of symbols, layout, or reading-aloud protocols. These elements can be particularly important for ensuring comprehension among primary-age students in LMIC contexts and may have influenced response behavior and data quality.

These limitations highlight the need for continued validation and refinement. Future work could include expert consultation across more regions, empirical validation of adapted items, incorporation of emerging domains, and evaluation of questionnaire design and administration procedures.

7.3 Implications for Questionnaire Design

The findings of this study have several implications for the design of context questionnaires to measure learning conditions in LMICs within large-scale comparative assessments at the end of primary education. While the results may also inform the development of context questionnaires for national assessments, the focus is on internationally comparable measurement in LMIC contexts, which was the aim of the study.

The implications presented below reflect the international and regional assessment frameworks and instruments included in the analysis while also identifying areas for future development and complementary extensions.

7.3.1 Coverage and Alignment with Conceptual Frameworks

One implication concerns the alignment between questionnaire content and the underlying integrated conceptual framework. The retained indicators were selected primarily based on their prevalence across assessment instruments, complemented by expert review. Items used to operationalize these indicators were drawn preferentially from established IEA content where available and formulated using simple and easily understood wording. Consequently, the selected items may not always capture the full breadth of the corresponding indicators, nor do they provide equal coverage of all elements of the project framework. Some areas are therefore

represented more extensively than others, reflecting both their prevalence across existing assessments and the selection approach adopted in this study. Users should assess whether the proposed item set provides sufficient coverage of the indicators and framework components relevant to their intended use. The framework–indicator mapping developed in this study can support this assessment and help identify areas where additional items may be needed.

7.3.2 Extension to Emerging and Evolving Domains

Another implication concerns the need to update questionnaire content as educational contexts evolve. While the present study focused on instruments developed up to approximately 2019, recent developments point to the increasing importance of digital learning environments and the use of generative artificial intelligence (AI) in education.

Future questionnaire designs may therefore consider incorporating selected items relating to digital learning environments and the use of AI. These could include school policies, teacher use, or student engagement with digital tools. Additional detail and illustrative item examples are discussed in the recommendation section.

7.3.3 Balancing Standardization and Contextual Adaptation

A central implication of the study is that some questionnaire content should remain adaptable to national contexts. While the proposed questionnaire item sets were designed to be broadly applicable across LMIC contexts, the findings indicate that certain constructs—particularly those related to socioeconomic conditions—may require additional flexibility. Even among LMIC-focused regional assessments, considerable variation was observed in the emphasis and granularity of SES-related indicators, such as resource deprivation and children’s involvement in income-generating and household activities.

In practice, this may involve supplementing the core questionnaire with country-specific modules, items, or response options to capture context-specific variation where needed.

7.3.4 Questionnaire Design Across Respondent Groups

The study results underscore the importance of carefully considering the respondent from whom information is collected. The proposed item set is based on an implementation model in which student, teacher, and school questionnaires are used, reflecting common operational constraints in LMIC contexts. When additional instruments are feasible, however, some constructs may be more appropriately allocated to other instruments. If a parent questionnaire can be administered, for example, constructs related to the student home environment and background may be more reliably obtained through parent reports than through student self-report. Within a cross-national assessment, respondent allocation should remain consistent across participating countries to ensure comparability, although the choice of respondent groups may differ between studies depending on operational feasibility.

7.3.5 Use of Optional and Context-Specific Items

The findings suggest that not all items in the proposed questionnaire sets need to be administered in every context. Some questions may have limited relevance in particular settings; items on language of instruction in predominantly monolingual contexts, for example, would have little utility. Where appropriate, optional or context-specific items can improve the efficiency and relevance of questionnaires while reducing respondent burden. Their use,

however, should be guided by clear criteria to ensure that key indicators remain comparable across countries.

7.4 Policy and Analytical Use

The proposed set of context indicators is designed not only to inform questionnaire development but also to support policy-relevant analysis and international monitoring in LMIC contexts. The following implications concern the interpretation and use of the resulting data.

7.4.1 Policy Relevance and Feasibility (NRC Perspective)

The qualitative review by NRCs from the LaNA 2023 Linking Study indicated that most proposed questionnaire items were perceived as relevant and feasible for use in their national contexts. This suggests that the selected indicators are broadly aligned with current policy priorities and information needs in LMICs (that participated in LaNA).

The review highlighted that feasibility and clarity are closely linked to contextual adaptation. Items that were simplified or generalized for cross-country applicability were, in some cases, flagged by experts as potentially problematic. These findings demonstrate the value of expert review in identifying issues that may not be evident from framework comparisons or empirical analyses alone.

7.4.2 Alignment with SDG Monitoring and Measurement Constraints

The findings also have implications for the use of ILSA context questionnaire data to support SDG 4 monitoring. Several SDG-relevant constructs traditionally captured through parent questionnaires or administrative data can, to some extent, be approximated through student questionnaires using simplified, age-appropriate formats. Evidence from regional assessments suggests that primary-age students can reliably report on selected aspects of SDG indicators, such as pre-primary participation, basic household conditions, and elements of the home learning environment. This is particularly relevant in LMIC contexts where parent questionnaires are often not feasible.

The analysis also highlighted important structural constraints. While some indicators can be operationalized through relatively simple questionnaire additions (e.g., grade repetition for deriving over-age status), others depend on the availability of additional variables, such as denominators or disaggregation measures, which may be collected through different instruments. In such cases, the corresponding items must be included, and the resulting data linked across sources. Indicators such as student–teacher qualification ratios, for example, require the integration of teacher and school-level data and cannot be derived from isolated questionnaire components.

7.4.3 Interpretation of Indicators for Monitoring Purposes

The empirical analyses conducted in this study, particularly the observed associations between context variables and achievement, should be interpreted in light of the standardized operationalization approach, which introduced certain constraints related to aggregation, data linkage, and cross-country comparability. These associations should therefore be used to inform the relevance and potential utility of the suggested questions. The proposed indicators are best suited to supporting system-level diagnostics, equity monitoring, and the identification of policy-relevant patterns, rather than explanatory modeling.

7.5 Recommendations

7.5.1 Use and Validation of the Proposed Questionnaire Item Sets

The results of this study support a differentiated approach to questionnaire development based on the developed suitability classification. This classification distinguishes between items for which prior evidence from existing assessments provides greater confidence regarding their likely functioning and items that require more extensive validation due to substantial modification and, hence, more limited transferable evidence.

A substantial share of items was classified as suitable, indicating that no major concerns were identified through expert review or empirical screening. These items have demonstrated acceptable performance in at least one context and can serve as a robust starting point for questionnaire development, although they should still undergo standard piloting procedures.

Items classified as “Consider Revision” or “Requires Validation and Revision” reflect cases where expert feedback or empirical evidence indicated potential issues related to clarity, differentiation, or response patterns. For these items, targeted refinement of wording, response formats, or contextual adaptation should be considered prior to use (see Section 7.5.2).

Items classified as “Requires Validation” are typically those that underwent more substantial adaptation or restructuring. While these items do not necessarily exhibit inherent weaknesses, their modified form lacks sufficient empirical evidence and therefore requires further testing to ensure reliable performance in the intended context.

7.5.2 Recommendations for Flagged Items

The following sections provide a more detailed review of questionnaire items that were classified as “Consider Revision,” “Requires Validation,” or “Requires Validation and Revision.”

Table 7-1 (student level) and Table 7-2 (teacher and school level) summarize the information underlying the final item suitability categorizations. The tables present the item identifier, indicator, and item stem, together with the ratings for policy relevance, clarity, differentiation, expert concerns, psychometric screening results, modification levels, and the resulting overall suitability classification. In addition, the tables summarize the identified underlying issues, specific NRC comments where available, and recommendations for how the respective items should be handled in future questionnaire development or adaptation processes.

For items marked for expert concern, the analysis draws on the qualitative feedback provided by NRCs from the LaNA 2023 Linking Study (see Section 6.7). For items flagged based on psychometric criteria (limited to items with minor or moderate adaptations), the evaluation is informed by the statistical properties of the corresponding variables, including indicators such as variability and association with achievement. None of the minimally or moderately adapted items exhibited high omission rates.

Being flagged does not imply that an item is unsuitable for use. The thresholds applied in this review were intentionally sensitive, to identify potential issues at an early stage. Some items may therefore still perform adequately in practice, particularly when adapted to specific contexts. This is especially relevant for items showing dominant response categories, which may nevertheless yield meaningful variation in different LMIC settings.

Table 7-1: Recommendation on Question Level (Students)

Level	ID	Indicator	Stem	Policy Relevance	Clarity	Differentiation	Expert Concern	Mean R ² (assoc.)	Mod. Level	Psych. Concern	Overall Suitability	ISSUE	RECOMMENDATION
STUDENT	5-A	Parents' literacy	How many adults in your home can read or write?	2	1.4	1.5	X	0.006	3		Validation and Consider Revision	Expert Clarity Ambiguous definition of "household" and "adult"; children in extended households unlikely to know literacy status of all adults	Option A: Clarify the definition of "adult" and the scope of the household. Option B: Follow PASEC approach — ask explicitly about different household members.
	6-A	Parents' Highest formal education	What is the highest level of education that your father (or male guardian) /mother (or ...	1.8	1.4	1.25	X	0.087	2		Consider Revision	Expert Clarity & Differentiation Students may not know parental education level; response categories rated too complex for this age group	Option A: Move item to parent questionnaire if administered. Option B: Reduce or simplify response categories. Note: Strong association with achievement in SEA-PLM.
	13-A	Type, material of house	What are the walls of your home mostly made of?	1.4	2	1.25	X	0.037	3		Validation and Consider Revision	Expert Differentiation no further reviewer comments provided	Option A: Deploy as optional item in low-resource contexts. Option B: Shift to parent questionnaire if administered; consider a more complex scale (e.g., MICS).
	14-A	Parental involvement in learning (incl. help with schoolwork/ Tuition)	How often does someone at home help you with your homework?	1.67	2	1	X	0.019	3		Validation and Consider Revision	Expert Differentiation no further reviewer comments provided	Option A: Revise response categories (e.g., include an "always" option). Option B: Add a filter question ("Do you get homework?") to improve clarity — note this increases item complexity.
	15-A	Assignment of Homework	About how often do you get homework? a) In mathematics b) In <language of test>/ reading	1.8	1.8	0.75	X	0.032	2		Consider Revision	Expert Differentiation	No specific recommendations.
	18-B	Nutrition, meals	How often do you have trouble paying attention in school because of these things? a) I ...	2	1.75	1.33	X	0.007	3		Validation and Consider Revision	Expert Differentiation no further reviewer comments provided	No specific recommendations.
	19-A	Physical and mental health (incl. Anxiety, Depression)	How often do you have trouble paying attention in school because of these things? a) I ...	2	1.75	1.33	X	0.015	3		Validation and Consider Revision	Expert Differentiation no further reviewer comments provided	No specific recommendations.
	26-A	Language of class instruction	How often does your reading/ language teacher use a language different from the <language...	2	1.8	1.25	X	0.009	3		Validation and Consider Revision	Expert Differentiation reviewers state that students typically have one teacher only	Option A: Adapt "reading/language teacher" as a country-specific term. Option B: Retain as optional item; relevant primarily in multilingual contexts.
	26-B	Language of class instruction	Which other language does your language teacher use the most when teaching?	1.75	1.75	1.25	X		3		Validation and Consider Revision	Expert Differentiation Reviewers state that "Other" category adds limited value; reviewers call for more context-specific language categories (language of instruction, mother tongue, learning environment)	Option A: Refocus scope — assess the extent to which language use is a hindrance to learning. Option B: Retain as optional item; relevant primarily in multilingual contexts.

Table 7-2: Recommendation on Question Level (Teachers and Schools)

Level	ID	Indicator	Stem	Policy Relevance	Clarity	Differentiation	Expert Concern	Mean R ² (assoc.)	Mod. Level	Psych. Concern	Overall Suitability	ISSUE	RECOMMENDATION
TEACHER	1-A	Teacher Gender	Which of these describes you?	2	1.75	1.75		0.022	2	X	Consider Revision	Psychometric Dominant category TIMSS: 86% female teachers	Retain for diagnostic purposes.
	5-A	Professional Teaching Qualification	Do you have a formal teaching qualification that qualifies you to teach in schools? (e....	2	2	1.83		0.014	2	X	Consider Revision	Psychometric Dominant category ERCE (PPIT14): 92% — binary degree variable lacks discriminatory power in LMIC contexts	Retain binary item for baseline comparability. Add modular follow-ups where greater differentiation is needed: 1.) Structural characteristics (institution type, duration, mode of study). 2.) Training content and practice (pedagogy, classroom practice, curriculum alignment). 3.) Perceived effectiveness (contribution to teaching performance).
	6-A	Subject-Matter Assignments (general /in sel. class)	Which of the following two subjects do you teach in the selected class? a) Mathematics ...	2	2	1.8		0.009	2	X	Consider Revision	Psychometric Dominant category ERCE (PPIT07_01/02/03): ~85%	Drop or retain for routing/verification purposes only. Note: very low association with achievement.
	10-B	Professional development (duration, quality)	If you answered yes to any of the activities above, to what extent did these formal pro...	2	2	1.4	X	0.002	2		Consider Revision	Expert Differentiation Country reviewers flag overlapping response categories ("Some" / "A little")	Consider revising the response scale (ERCE uses a 1–5 rating scale). Note: low association with achievement.
SCHOOL	20-A	Availability of textbooks	How does your school make textbooks available to students in <target grade>? a) Mathema...	2	2	2		0.026	2	X	Consider Revision	Psychometric Dominant category	Retain to identify low-resourced student groups across countries/regions. Consider considering e-textbooks in future cycles.
	22-A	Digital learning resources	How many digital devices does your school have for use by <target grade> students?	1.75	1.5	1.33	X	0.027	3		Validation and Consider Revision	Expert Clarity Conceptual mismatch	Option A: Revise wording to school-level access — e.g., "Approximately how many digital devices (e.g., computers, tablets) are available at your school for students to use?" Option 2: Shift focus from device count to availability/access.
	25-A	Shifts	Does your school have double shifts or double sessions? (one group of students attends ...	2	1.6	1.75		0.005	1	X	Consider Revision	Psychometric Dominant category	Consider making shifts a country-specific term (countries may call it "double-flow", etc,
	36-A	Nutrition, meals provided by school	Does your school provide free meals for students? a) Breakfast b) Lunch	2	2	2		0.019	1	X	Consider Revision	Psychometric Dominant category	Retain for diagnostic purposes.

7.5.2.1 *Student Level*

The review of the student-level items indicates that most concerns relate to expert-identified issues of clarity and differentiation in the expected response patterns rather than the underlying policy relevance of the constructs. Across the flagged items, three main patterns emerge.

First, some items raise concerns related to clarity and respondent burden. Students at the primary level may not have reliable knowledge of parental characteristics, and key concepts (e.g., “adult,” “household,” or levels of education) may be interpreted inconsistently. In addition, response formats are sometimes too complex for the target age group. These issues suggest the need for clearer definitions, simplified response categories, and, where feasible, the relocation of such items to a parent questionnaire or the use of structured proxy approaches.

Second, limited differentiation is a recurring issue across a broad range of items, including those related to household conditions, parental involvement, homework practices, student well-being, and language use in the classroom. In these cases, response distributions are expected to show low variability or reflect poorly differentiated or overlapping response categories. Addressing this requires careful revision of response scales to ensure sufficient spread and consideration of alternative formulations where appropriate.

Finally, some items are highly context-dependent. Indicators such as housing materials are primarily informative in low-resource settings, while items on language use are most relevant in multilingual education systems. For these items, country-specific adaptations, context-sensitive response options, or optional implementation may be more appropriate than inclusion as core indicators. Where relevant, the underlying construct may also be refocused (e.g., by capturing the extent to which language use constitutes a barrier to learning).

7.5.2.2 *Teacher and School Level*

The review of the teacher- and school-level items indicates that most concerns relate to psychometric limitations, particularly dominant response categories, rather than issues of clarity or respondent burden. As a result, many items show limited variability and reduced analytical power.

For teacher-level items, core variables such as gender, teaching qualification, and subject assignment should generally be retained for descriptive and diagnostic purposes. However, where binary or highly skewed measures are used (e.g., teaching qualification), their analytical value can be strengthened by adding items capturing training structure, content, and perceived effectiveness. Items with very low variability, such as subject assignment in primary contexts, are best treated as routing or verification items rather than substantive indicators. Where expert concerns were raised (e.g., professional development), revising response scales to avoid overlapping categories is recommended.

At the school level, items on resources and school conditions (e.g., textbooks, shifts, school meals) also frequently show dominant categories. These should be retained to identify disadvantaged or low-resource contexts, even if their explanatory power is limited. In contrast, the item on digital learning resources requires conceptual revision, as the current formulation does not align with how resources are organized in schools. A shift toward measuring school-level availability or access rather than grade-specific counts is recommended.

7.5.3 Recommendations for SDG-Oriented Measurement

The proposed questionnaire item sets were designed to balance feasibility, respondent capacity, and cross-country comparability. Questionnaire design may differ, however, when the primary objective is the measurement of SDG-related indicators.

Recent work illustrates how contextual, input, and process indicators drawn from ILSAs and complementary data sources can support the monitoring and projection of progress toward SDG 4.1.1, while also highlighting persistent data gaps for many LMICs (Kaplan, Jude, Harra, & Stampka, 2025). For indicators related to household characteristics, parental background, and early childhood experiences, a more comprehensive measurement approach may be required to fully align with SDG reporting standards. This may involve more detailed item formulations, extended response categories, or the use of additional instruments. In particular, the administration of a dedicated home or parent questionnaire may be advisable to collect such information more comprehensively and reliably. Existing resources such as *Analysis of Context Questionnaires in ILSAs* (UIS, 2025) may provide useful guidance for aligning questionnaire content with SDG indicator definitions and reporting requirements.

In addition to items directly measuring SDG-related constructs, questionnaire design should also consider the inclusion of necessary disaggregation variables and prerequisites, as outlined in Section 6.3.3.2 (see Table 6-13). These may be required across multiple indicators, for example as variables for disaggregation, as inputs for the construction of composite measures such as wealth indices, or as denominators for derived indicators (e.g., pupil–trained teacher ratios). Given that some SDG indicators are still under development and are likely to be revised in the post-2030 framework, users should ensure that questionnaire content reflects the most recent methodological updates and reporting requirements.

7.5.4 Recommendations Based on Expert-Identified Indicator Gaps

Country experts also identified a limited number of areas where more comprehensive information could be collected, reflecting context-specific priorities across LMIC settings. At the student level, these relate mainly to learning resources and constraints (e.g., access to materials, ICT, and child labor). At the teacher level, suggestions focus on pedagogical processes and professional conditions, including instructional support, formative practices, and working conditions. At the school level, proposed additions concern infrastructure, environmental conditions, and governance (e.g., water and sanitation, physical conditions, multigrade teaching, and management structures). These inputs primarily relate to contextual and process-related dimensions and are provided as targeted extensions for further consideration in specific contexts. Further details are provided in Appendix A4.

7.5.5 Recommendations from Framework–Indicator Mapping

The comparison of retained and non-retained indicators across the integrated conceptual framework provides further insight into the overall balance of the proposed questionnaire item sets. Of the 184 indicators considered, 90 were retained and 94 were not retained, with notable differences across framework dimensions.

The retained indicators are more strongly concentrated in antecedent domains, particularly at the school and teacher level, where structural characteristics and background conditions are well represented. Indicators related to contextual and process-oriented dimensions, especially at the school and student level, were less frequently retained. Certain content-related

dimensions, particularly those related to implemented and experienced processes, also show comparatively lower coverage.

While the proposed questionnaire item sets, selected based on their representation in a broader range of the questionnaires, provide a robust core for describing structural conditions, they may provide less comprehensive coverage of dynamic aspects of teaching, learning environments, and classroom processes. Users seeking more comprehensive coverage of learning conditions should consider incorporating selected indicators from the non-retained set, particularly those related to instructional processes, classroom practices, and student experiences.

7.5.6 Practical Considerations for Questionnaire Design

The proposed questionnaire item sets combine sensitivity to low-resource contexts with alignment to established international assessment practices. They include selected indicators of material living and learning conditions alongside standardized constructs related to teaching and school climate. Although they extend the contextual coverage of the existing LaNA questionnaires, their simplified wording and response formats, limited skip logic, and careful selection of items aim to maintain feasibility and accessibility, particularly for primary-level respondents.

7.6 Areas for Further Development and Outlook

7.6.1 Emerging Domains and Strengthening the Evidence Base

Educational systems are changing rapidly, particularly with respect to digitalization and the role of artificial intelligence in teaching and learning. Frameworks such as TIMSS 2027 (Davier & Kennedy, 2025) and the upcoming Media and AI Literacy (MAIL) framework in PISA 2029 cycle (OECD, 2026) are reflecting these developments across the ILSA landscape. Future questionnaire designs focusing on LMIC contexts may also need to incorporate aspects related to access to digital tools, their pedagogical use, and emerging forms of student engagement with technology.

The uneven pace of digital transformation across LMICs reinforces the need for adaptive, modular frameworks that can accommodate diverse contexts. The evidence base for the proposed indicators also requires further strengthening across additional LMIC regions, including testing item performance across countries, refining measures with low variability or limited differentiation, and examining indicator robustness for policy-relevant analysis and monitoring.

7.6.2 Administration Procedures, Timing, and Visual Design

While the present study focused on item content and structure, administration procedures, completion time, and visual design also play an important role in ensuring accessibility for primary-age respondents. A 30-minute administration period, often used as a threshold for administering student questionnaires, may provide a useful reference point but should not be treated as a fixed limit independent of questionnaire content and administration mode. Regional assessments demonstrate that substantially broader contextual information can be collected from primary-age students. In PASEC 2019, which included the smallest number of student background variables among the regional assessments reviewed, approximately 100 variables were administered within a collective questionnaire session of about 45 minutes. SACMEQ IV collected close to 300 student-level variables. These examples suggest that the

feasible scope of a questionnaire depends not only on its duration, but also on procedural arrangements, including collective administration with questions read aloud where appropriate, as implemented in PASEC.

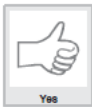
Questionnaire development should nevertheless continue to prioritize simple wording, accessible response formats, and manageable cognitive demands to keep respondent burden low. Completion time, respondent burden, and item omission rates should be monitored and evaluated empirically during piloting. Rotated questionnaire forms may also be considered, where broader content coverage is required without increasing the burden on individual respondents.

Visual design features can support students in understanding both the content of questions and the expected response format. In particular, the use of icons, visual cues, and clearly structured response formats can facilitate comprehension. For example, pictorial representations of response options, visual anchors for frequency scales, or differentiated checkmark formats may help guide students through the questionnaire and reduce ambiguity in how answers should be marked. Examples from ERCE and TIMSS student questionnaires illustrate how such elements can support comprehension, especially for younger students or those with limited exposure to formal assessment settings (see Figure 7-1 and Figure 7-2). These features may also reduce cognitive burden by making questions more intuitive and visually navigable. Such design elements may be particularly beneficial in LMIC contexts, where students may have varying levels of familiarity with written questionnaires and standardized response formats.

15. **What do you think of the following statements about mathematics?**

For each of the following sentences, mark with an X only one answer option (Yes, More or less, No).

15.1. In general, I do well in mathematics.



Yes



More or less



No

Figure 7-1: Visualization example from ERCE Grade 3 Questionnaire

Mathematics in School

1. How much do you agree with these statements about mathematics?

*Click **one** circle for each row.*

	☒ Agree a lot	☐ Agree a little	☐ Disagree a little	☐ Disagree a lot
a) Mathematics is easy for me	☐	☐	☐	☐
b) I like to solve mathematics problems	☐	☐	☐	☐

Figure 7-2: Visualization example from TIMSS Grade 4 Questionnaire

Questionnaire administration procedures are likewise critical for ensuring consistent and reliable data collection. PASEC and ERCE emphasize the use of scripted instructions, the clear and slow reading of questions, and limitations on providing additional explanations. In contexts where student reading proficiency is limited, reading questionnaire items aloud while

maintaining standardized administration procedures may improve accessibility. Administrators should support students by clarifying procedural aspects, building confidence, and ensuring that all items are attempted without influencing responses.

These considerations highlight that effective questionnaire implementation depends not only on item design, but also on visual presentation and administration procedures. Where feasible, pretesting and cognitive interviewing with students can help identify potential misunderstandings and refine item wording and response formats.

8 Appendices

8.1 Appendix A1: Suggested Student Questionnaire Items

1. Which of these describes you?

[Indicator: Gender]

- ☐ Girl
- ☐ Boy
- ☐ <Other>

2. How old are you?

[Indicator: Age]

___ years old

3A. How often do you speak the <language of the test> at home?

[Indicator: Language experience]

- ☐ Always
- ☐ Almost always
- ☐ Sometimes
- ☐ Never

3B. What language do you speak at home most of the time?

[Indicator: Language experience]

- ☐ <language of test>
- ☐ <language 2>
- ☐ <language 3>
- ☐ <Other> (Note: If two languages are spoken at the same frequency, choose the one you learned first.)

4. Where were you born?

[Indicator: Immigration Status]

- ☐ In this same village/town/city where I live now
- ☐ In another village/town/city in <this country>
- ☐ In another country
- ☐ I don't know

5. How many adults in your home can read or write?

[Indicator: Parents' literacy]

- ☐ No adults at home can read or write
- ☐ One adult at home can read or write
- ☐ More than one adult at home can read or write
- ☐ I don't know

6. What is the highest level of education that your father (or male guardian) /mother (or female guardian) have completed?

[Indicator: Parents' Highest formal education]

	Mother or female caregiver	Father or male caregiver
a) Did not go to school	☐	☐
b) <Primary education—ISCED Level 1>	☐	☐
c) <Lower secondary education—ISCED Level 2>	☐	☐
d) <Upper secondary education—ISCED Level 3>	☐	☐
e) <Post-secondary non-tertiary education or short-cycle tertiary education—ISCED Level 4 or 5>	☐	☐
f) <Bachelor's or equivalent level—ISCED Level 6>	☐	☐
g) <Master's or equivalent or doctoral or equivalent level—ISCED Level 7 or 8>	☐	☐
g) I don't know	☐	☐
h) Not applicable	☐	☐

7. Did you go to a preschool or kindergarten before starting primary school?

[Indicator: Participation in formal pre-primary education]

- ☐ Yes, for 2 years or more
- ☐ Yes, for between 1 year and 2 years
- ☐ Yes, for less than 1 year
- ☐ No
- ☐ I don't know

8. Have you ever repeated a grade?

[Indicator: Grade repetition]

- ☐ Yes
- ☐ No

9A. About how often are you absent from school?

[Indicator: Absenteeism]

- ☐ Never or almost never
- ☐ About once every two months
- ☐ About once per month
- ☐ About once every two weeks
- ☐ About once a week or more

9B. About how often do you arrive late to school?

[Indicator: Absenteeism]

- ☐ Never or almost never
- ☐ About once every two months
- ☐ About once per month
- ☐ About once every two weeks
- ☐ About once a week or more

10A. How many people currently live in your home? (Including yourself.)

[Indicator: Household composition (number of people living at home)]

- ☐ 1–3
- ☐ 4–6
- ☐ 7–9

☐ 10 or more

10B. How many siblings (brothers and sisters) do you have?

[Indicator: Household composition (number of people living at home)]

☐ None

☐ 1

☐ 2

☐ 3

☐ 4

☐ More than 4

11. How many books are there in your home? (Do not count schoolbooks, newspapers, magazines)

[Indicator: Number of books in the home]

☐ No books

☐ 1–10

☐ 11–20

☐ 21–30

☐ 31–50

☐ 51–100

☐ More than 100

12. Do you have any of these things at your home?

[Indicator: Household possessions and services]

	Yes	No
a) Running water	☐	☐
b) Electricity	☐	☐
c) Refrigerator (fridge)	☐	☐
d) Mobile phone	☐	☐
e) Computer or tablet	☐	☐
f) Access to the internet	☐	☐
g) Study desk/table for your use	☐	☐
h) Your own <bedroom>	☐	☐
i) Television	☐	☐
j) Stove for cooking food	☐	☐
k) Car or motorcycle	☐	☐
l) Flush toilet	☐	☐
m) <country-specific>	☐	☐

13. What are the walls of your home mostly made of?

[Indicator: Type, material of house]

☐ Cardboard

☐ Grass

☐ Mud and sticks

☐ Stones

☐ Metal sheets

☐ Wood

☐ Cut stones or bricks

☐ Other

14. How often does someone at home help you with your homework?

[Indicator: Parental involvement in learning (incl. help with schoolwork/ Tuition)]

☐ I do not get any homework

☐ Never

☐ Sometimes

☐ Most of the time

15. About how often do you get homework?

[Indicator: Assignment of Homework]

	I do not get any homework	Once or twice a month	Once or twice a week	Most days of the week
a) In mathematics	☐	☐	☐	☐
b) In <language of test>/ reading	☐	☐	☐	☐

16A. How much do you agree with these statements about reading?

[Indicator: Enjoyment of reading/ mathematics]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I like talking about what I read with other people	☐	☐	☐	☐
b) I would like to have more time for reading	☐	☐	☐	☐
c) I enjoy reading	☐	☐	☐	☐
d) I like to read things that make me think	☐	☐	☐	☐

16B. How much do you agree with these statements about learning mathematics?

[Indicator: Enjoyment of reading/ mathematics]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I enjoy learning mathematics	☐	☐	☐	☐
b) I like mathematics	☐	☐	☐	☐
c) I look forward to mathematics lessons	☐	☐	☐	☐
d) Mathematics is one of my favorite subjects	☐	☐	☐	☐

17A. How much do you agree with these statements about reading?

[Indicator: Confidence in reading/ mathematics]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I usually do well in reading	☐	☐	☐	☐
b) Reading is easy for me	☐	☐	☐	☐
c) I understand most of what I read	☐	☐	☐	☐

17B. How much do you agree with these statements about mathematics?

[Indicator: Confidence in reading/ mathematics]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
--	-------------	-------------------	----------------------	-------------------

a) I usually do well in mathematics	☐	☐	☐	☐
b) Mathematics is easy for me	☐	☐	☐	☐
c) I am good at working out difficult mathematics problems	☐	☐	☐	☐

18A. Which meals do you eat on a normal school day?

[Indicator: Nutrition, meals]

	Yes	No
a) Breakfast	☐	☐
b) Lunch	☐	☐
c) Dinner	☐	☐

18B. How often do you have trouble paying attention in school because of these things?

[Indicator: Nutrition, meals]

	Every day	Almost every day	Sometimes	Never
a) I feel hungry	☐	☐	☐	☐

19A. How often do you have trouble paying attention in school because of these things?

[Indicator: Physical and mental health (incl. Anxiety, Depression)]

	Every day	Almost every day	Sometimes	Never
a) I feel tired	☐	☐	☐	☐

19B. Do you agree with these statements about your health:

[Indicator: Physical and mental health (incl. Anxiety, Depression)]

	Yes	No
a) I can clearly see the words the teacher writes on the board	☐	☐
b) I can clearly hear the teacher when he or she is speaking	☐	☐
c) I have a physical impairment that makes it hard for me to walk or go up and down stairs	☐	☐
d) I have a physical impairment that makes it hard for me to hold or use small things like a pencil or scissors	☐	☐
e) I get sick so often that sometimes I cannot play, do things, or go to school	☐	☐

20. How much do you agree with these statements about your school?

[Indicator: Well being at school]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I like being in school	☐	☐	☐	☐
b) I feel safe when I am at school	☐	☐	☐	☐
c) I feel like I belong at this school	☐	☐	☐	☐

21. How often do these things happen in your class?

[Indicator: Classroom climate]

	Every or almost every day	Once or twice a week	Once or twice a month	Never or almost never
a) Students don't listen to what the teacher says	☐	☐	☐	☐
b) There is too much noise for students to work well	☐	☐	☐	☐
c) My teacher has to wait a long time for students to be quiet	☐	☐	☐	☐
d) Students interrupt the teacher	☐	☐	☐	☐
e) My teacher has to keep telling us to follow the classroom rules	☐	☐	☐	☐

22. During this year, how often have other students from your school done any of the following things to you (in person, through messaging, or through social media)?

[Indicator: Experience of bullying/ harassment]

	At least once a week	Once or twice a month	A few times	Never
a) Made fun of me or called me names	☐	☐	☐	☐
b) Left me out of their games or activities	☐	☐	☐	☐
c) Spread lies about me	☐	☐	☐	☐
d) Stole something from me	☐	☐	☐	☐
e) Damaged something of mine on purpose	☐	☐	☐	☐
f) Hit or hurt me	☐	☐	☐	☐
g) Made me do things I didn't want to do	☐	☐	☐	☐
h) Sent me nasty or hurtful messages online	☐	☐	☐	☐
i) Shared nasty or hurtful information about me online	☐	☐	☐	☐
j) Threatened me	☐	☐	☐	☐

23. How often do you borrow books (including ebooks) from your <class/ school or local library>?

[Indicator: Access to books from community/school/class library]

- ☐ I do not have access to a <class/ school or local library>
- ☐ At least once a week
- ☐ Once or twice a month
- ☐ A few times a year
- ☐ Never or almost never

24. How often do you do these things outside of school?

[Indicator: House chores/ Work]

	Every day	Almost every day	Sometimes	Never
a) Do house chores (e.g., cleaning, cooking, fetching water or firewood)	☐	☐	☐	☐
b) Do farm work (e.g., planting, harvesting, caring for animals)	☐	☐	☐	☐
c) Help to sell things or earn money (e.g., selling at a market, helping in a shop)	☐	☐	☐	☐
d) Take care of family members (e.g., looking after siblings or sick relatives)	☐	☐	☐	☐

25. How long does it take you to get to school each day?

[Indicator: Commuting to school]

- ☐ Less than 30 minutes
- ☐ 30 minutes up to 1 hour
- ☐ 1 hour to 2 hours
- ☐ 2 hours or more

26A. How often does your reading/ language teacher use a language different from the <language of test> when teaching?

[Indicator: Language of class instruction]

- ☐ Never
- ☐ Sometimes
- ☐ Often

26B. Which other language does your language teacher use the most when teaching?

[Indicator: Language of class instruction]

- ☐ None, my teacher only uses the <language of test>
- ☐ <language 1>
- ☐ <language 2>
- ☐ <language 3>
- ☐ <Other>

27A. How much do you agree with these statements about your reading lessons?

[Indicator: Perception of Instruction]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I know what my teacher expects me to learn in each lesson	☐	☐	☐	☐
b) My teacher is easy to understand	☐	☐	☐	☐
c) I am interested in what my teacher says	☐	☐	☐	☐
d) My teacher asks me to explain my answers	☐	☐	☐	☐
e) My teacher has clear answers to my questions	☐	☐	☐	☐

27B. How much do you agree with these statements about your mathematics lessons?

[Indicator: Perception of Instruction]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I know what my teacher expects me to learn in each session	☐	☐	☐	☐
b) My teacher is easy to understand	☐	☐	☐	☐
c) I am interested in what my teacher says	☐	☐	☐	☐
d) My teacher asks me to explain my answers	☐	☐	☐	☐
e) My teacher is good at explaining mathematics	☐	☐	☐	☐

8.2 Appendix A2: Suggested Teacher Questionnaire Items

1. Which of these describes you?

[Indicator: Teacher Gender]

- ☐ Female
- ☐ Male
- ☐ <Other>

2. How old are you?

[Indicator: Teacher Age]

- ☐ Under 25
- ☐ 25–29
- ☐ 30–39
- ☐ 40–49
- ☐ 50–59
- ☐ 60 or more

3A. By the end of this school year, how many years will you have been teaching altogether?

[Indicator: Years of Teaching]

_____ years (Please round to the nearest whole number.)

3B. By the end of this school year, how many years will you have been teaching at this school?

[Indicator: Years of Teaching]

_____ years (Please round to the nearest whole number.)

4. What is the highest level of formal education you have completed?

[Indicator: Highest level of education]

- ☐ Did not complete <Upper secondary education—ISCED Level 3>
- ☐ <Upper secondary education—ISCED Level 3>
- ☐ <Post-secondary non-tertiary education—ISCED Level 4>
- ☐ <Short-cycle tertiary education—ISCED Level 5>
- ☐ <Bachelor's or equivalent Level—ISCED Level 6>
- ☐ <Master's or equivalent level—ISCED Level 7>
- ☐ <Doctoral or equivalent level—ISCED Level 8>

5. Do you have a formal teaching qualification that qualifies you to teach in schools? (e.g., a teacher training certificate, diploma, or degree, completed either as part of or in addition to your highest level of education)

[Indicator: School degree]

- ☐ Yes
- ☐ No

6. Which of the following two subjects do you teach in the selected class?

[Indicator: Subjects taught by teacher (general /in sel. Class)]

	Yes	No
a) Mathematics	☐	☐
b) <language of test>	☐	☐

7. Which <grade(s) or class level(s)> do you currently teach in this school?

[Indicator: Grades/ Classes being taught]

- ☐ <Grade or class level 1>
☐ <Grade or class level 2>
☐ <Grade or class level 3>
☐ <Grade or class level 4>
☐ <Grade or class level 5>
☐ <Grade or class level 6>
☐ <Other>

8. Did you complete any pre-service teacher training? (This may have been completed as part of a regular degree or as separate teacher training).

[Indicator: Teacher Preparation - general (duration/quality/Experiences)]

- ☐ No
☐ Yes, for up to 1 year
☐ Yes, for more than 1 and up to 2 years
☐ Yes, for more than 2 and up to 3 years
☐ Yes, for more than 3 years

9A. During your <post-secondary> education, what was your major or main area(s) of study? (If you did not complete any <post-secondary> education, mark “No <post-secondary> education”).

[Indicator: Area(s) of Post-secondary Study / Teacher]

	Yes	No
a) No <post-secondary> education	☐	☐
b) Education—Primary/Elementary	☐	☐
c) Education—Secondary	☐	☐
d) Mathematics	☐	☐
e) <language of test>	☐	☐
f) Other	☐	☐

9B. If your major or main area of study was education, did you have a specialization in any of the following?

[Indicator: Area(s) of Post-secondary Study / Teacher]

	Yes	No
a) Mathematics	☐	☐
b) <language of test>	☐	☐
c) Other subject	☐	☐

10A. In the past two years, have you participated in formal professional development in the following areas, and how important are these areas for your current teaching position?

[Indicator: Professional development (duration, quality)]

	Participated: Yes	Participated: No	Importance: Not important	Importance: Important	Importance: Very important

a) Teaching reading comprehension skills or strategies	☐	☐	☐	☐	☐
b) Integrating reading literacy across the curriculum	☐	☐	☐	☐	☐
c) Teaching mathematics content and problem-solving	☐	☐	☐	☐	☐
d) Pedagogical methods for teaching reading or mathematics	☐	☐	☐	☐	☐
e) Integrating technology into instruction	☐	☐	☐	☐	☐
f) Assessing students' learning in reading or mathematics	☐	☐	☐	☐	☐
g) Addressing students' diverse language, learning, or inclusion needs	☐	☐	☐	☐	☐
h) Differentiating instruction based on student levels and interests	☐	☐	☐	☐	☐

10B. If you answered yes to any of the activities above, to what extent did these formal professional development activities improve your teaching practices?

[Indicator: Professional development (duration, quality)]

- ☐ A lot
☐ Some
☐ A little
☐ Not at all

11. How confident are you in doing each of these things when teaching this class?

[Indicator: Confidence in teaching (methods)]

	Very high	High	Medium	Low
a) Having my students develop critical thinking skills (e.g., debate, evaluate different opinions and substantiate their own)	☐	☐	☐	☐
b) Posing interesting or challenging questions that make my students think	☐	☐	☐	☐
c) Enhancing students' social-emotional skills (e.g., empathy, self-management, tolerance)	☐	☐	☐	☐
d) Explaining difficult content in several different ways (e.g., using a variety of examples, metaphors, diagrams, models, or instructional materials)	☐	☐	☐	☐
e) Using varied teaching strategies and alternating among them (e.g., discussions, projects, dramatization, problem solving, etc.)	☐	☐	☐	☐

12. How often do you have these feelings about being a teacher?

[Indicator: Job satisfaction (incl. General working conditions)]

	Very often	Often	Sometimes	Never
a) I am content with my profession as a teacher	☐	☐	☐	☐
b) I find my work full of meaning and purpose	☐	☐	☐	☐
c) I am enthusiastic about my job	☐	☐	☐	☐
d) My work inspires me	☐	☐	☐	☐
e) I am proud of the work I do	☐	☐	☐	☐
f) I feel appreciated as a teacher	☐	☐	☐	☐
g) I enjoy the challenges of teaching	☐	☐	☐	☐

13. During the last month, did any of the following situations prevent you from going to work?

[Indicator: Teacher Absence]

	Yes,	No
a) Health-related reasons (e.g., illness, mental or emotional health problems, medical appointments)	☐	☐
b) Family-related reasons (e.g., caring for family members, family emergency)	☐	☐
c) Workplace or labor-related reasons (e.g., conflicts or violence at school, strike)	☐	☐
d) Transport, weather, or other external reasons (e.g., lack of transport, extreme weather or hazards)	☐	☐

14. How long is your employment contract with this school?

[Indicator: Type of contract with school/ Position]

- ☐ For an indefinite term (tenured or appointed position)
- ☐ For a term of one year or more with the possibility of continuing (interim, contract)
- ☐ For a term of less than one year (substitutes, etc.)
- ☐ Other

15. In addition to your regular teaching at this school, how many hours per week do you spend on other paid work (including private tutoring or any other job)?

[Indicator: Additional paid jobs]

- ☐ None
- ☐ Fewer than 10 hours per week
- ☐ 10 to 19 hours per week
- ☐ 20 to 30 hours per week
- ☐ More than 30 hours per week

16A. Which of the following items do you have in your classroom?

[Indicator: Teaching/ learning resources and Infrastructure]

	Yes	No
a) Teacher's guide	☐	☐
b) Electricity/power	☐	☐
c) Working computer available for teaching	☐	☐
d) Interactive whiteboard/ smartboard	☐	☐
e) Internet available for teaching	☐	☐
f) Wall chart of any kind	☐	☐
g) Class set of student basic supplies (pencil/pen and exercise book)	☐	☐
h) Sufficient chairs or benches	☐	☐
i) Sufficient desks or tables	☐	☐

16B. How does your school make textbooks available to students in this class?

[Indicator: Teaching/ learning resources and Infrastructure]

	There are no textbooks	One textbook is shared	One textbook is shared	Most students have	Each student has their

	for students	by more than two students	by two students	their own textbook, but some have to share	own textbook
a) Mathematics	☐	☐	☐	☐	☐
b) <language of test>	☐	☐	☐	☐	☐

17A. How many books with different titles are in your classroom library or reading corner?

[Indicator: School- or Classroom library and use]

- ☐ My classroom does not have a library or reading corner
- ☐ Fewer than 25
- ☐ 26 to 50
- ☐ 51 to 100
- ☐ More than 100

17B. Can the students borrow books from the classroom library or reading corner to take home?

[Indicator: School- or Classroom library and use]

- ☐ Yes
- ☐ No

18. How would you characterize each of the following within your school?

[Indicator: EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)]

	Very high	High	Medium	Low	Very Low
a) Teachers' understanding of the school's curricular goals	☐	☐	☐	☐	☐
b) Teachers' degree of success in implementing the school's curriculum	☐	☐	☐	☐	☐
c) Teachers' expectations for student achievement	☐	☐	☐	☐	☐
d) Teachers' ability to inspire students	☐	☐	☐	☐	☐

19. How would you characterize each of the following within your school?

[Indicator: EAS_Parental Support]

	Very high	High	Medium	Low	Very Low
a) Parental involvement in school activities	☐	☐	☐	☐	☐
b) Parental commitment to ensure that students are ready to learn	☐	☐	☐	☐	☐
c) Parental expectations for student achievement	☐	☐	☐	☐	☐
d) Parental support for student achievement	☐	☐	☐	☐	☐

20A. How often have these things happened at your school in the past year?

[Indicator: SOS_Personal Safety (including bullying)]

	At least once a week	Once or twice a month	A few times	Never

a) A student threatened another student	☐	☐	☐	☐
b) A student hit or hurt another student	☐	☐	☐	☐
c) A student was discriminated against/excluded by other students because of his/her appearance, ethnicity, or social class	☐	☐	☐	☐
d) A student insulted a teacher	☐	☐	☐	☐

20B. How much do you agree or disagree with the following statements about your current school?

[Indicator: SOS_Personal Safety (including bullying)]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) I feel safe at this school	☐	☐	☐	☐

21A. How much do you agree or disagree with the following statements about your current school?

[Indicator: SOS_Student Behavior & Respect]

	Agree a lot	Agree a little	Disagree a little	Disagree a lot
a) This school's security policies and practices are sufficient	☐	☐	☐	☐
b) The students behave in an orderly manner	☐	☐	☐	☐
c) The students are respectful of the teachers	☐	☐	☐	☐
d) The students respect school property	☐	☐	☐	☐
e) This school has clear rules about student conduct	☐	☐	☐	☐
f) This school's rules are enforced in a fair and consistent manner	☐	☐	☐	☐

21B. How often do the following situations occur in the <classroom assessed>?

[Indicator: SOS_Student Behavior & Respect]

	Very often	Often	Sometimes	Never
a) When the class starts, I have to wait a long time for the students to be silent	☐	☐	☐	☐
b) I lose a lot of time because students disrupt the class	☐	☐	☐	☐
c) Students make a lot of noise in the classroom	☐	☐	☐	☐

22. Since the beginning of this school year, how often have the following people visited your classroom or discussed your teaching with you (e.g., through observation, feedback, or pedagogical meetings)?

[Indicator: External Supervision & Inspection]

	At least once a week	Once or twice a month	A few times a year	Never
a) External supervisors (e.g., school inspector, educational adviser, pedagogical advisor)	☐	☐	☐	☐
b) School management (e.g., principal, deputy principal, pedagogical coordinator)	☐	☐	☐	☐

c) Internal pedagogical staff (e.g., mentor, senior teacher, subject coordinator)	☐	☐	☐	☐
---	--------------------------	--------------------------	--------------------------	--------------------------

23. How many students are in this class?

[Indicator: Class size/ number of students on day of test]

	Male students (Write in the number.)	Female students (Write in the number.)
a) How many students are in this class in total?	☐	☐
b) How many of these students are in the <target grade>?	☐	☐

24A. In your view, to what extent do the following limit how you teach this class?

[Indicator: Teaching conditions/ Teaching Constraints]

	Not at all	Some	A lot
a) Students lacking prerequisite knowledge or skills	☐	☐	☐
b) Students suffering from lack of basic nutrition	☐	☐	☐
c) Students suffering from not enough sleep	☐	☐	☐
d) Students absent from class	☐	☐	☐
e) Disruptive students	☐	☐	☐
f) Uninterested students	☐	☐	☐
g) Distracted students	☐	☐	☐
h) Students with mental, emotional, or psychological impairment	☐	☐	☐
i) Students with difficulties understanding the language of instruction	☐	☐	☐

24B. To what extent do the following limit how you teach this class?

[Indicator: Teaching conditions/ Teaching Constraints]

	Not at all	Some	A lot
a) Lack of basic resources or equipment (e.g., textbooks, furniture, electricity, or computers and internet access)	☐	☐	☐
b) Poor school infrastructure	☐	☐	☐
c) Safety or security concerns in or around the school	☐	☐	☐
d) Teaching several grade levels in the same class	☐	☐	☐

25. What languages do you use when teaching your students in this class?

[Indicator: Language of instruction]

- ☐ Only <language of instruction>
- ☐ Mostly <language of instruction> but sometimes students' home languages
- ☐ Sometimes <language of instruction> and sometimes students' home languages
- ☐ Mostly students' home languages
- ☐ Always students' home languages

26A. In a typical week, how much time do you spend teaching mathematics to the students in this class?

[Indicator: Instructional time (Language/ Reading or mathematics)]

_____ minutes per week (Please convert the number of hours into minutes.)

26B. In a typical week, how much time do you spend on <language of test> language instruction and/or activities with the students?

[Indicator: Instructional time (Language/ Reading or mathematics)]

_____ minutes per week

27. How often do you do the following in teaching this class?

[Indicator: Instructional activities (general, Language/ Reading, Math)]

	Every or almost every lesson	About half the lessons	Some lessons	Never
a) Relate the lesson to students' daily lives	☐	☐	☐	☐
b) Ask students to explain their answers	☐	☐	☐	☐
c) Communicate goals or objectives for the lesson to the students	☐	☐	☐	☐
d) Ask students to complete challenging exercises that require them to go beyond the instruction	☐	☐	☐	☐
e) Encourage classroom discussions among students	☐	☐	☐	☐
f) Link new content to students' prior knowledge	☐	☐	☐	☐
g) Ask students to decide their own problem solving procedures	☐	☐	☐	☐
h) Check or review students' homework or classwork	☐	☐	☐	☐
i) Take a written quiz or test	☐	☐	☐	☐
j) Perform differentiated activities according to abilities	☐	☐	☐	☐
k) Have students memorize or repeat important facts, rules, or examples	☐	☐	☐	☐

8.3 Appendix A3: Suggested School Questionnaire Items

1. Which of these describes you?

[Indicator: Gender]

- ☐ Female
☐ Male
☐ <Other>

2. How old are you?

[Indicator: Age]

- ☐ Under 25
☐ 25–29
☐ 30–39
☐ 40–49
☐ 50–59
☐ 60 or more

3A. By the end of this school year, how many years will you have been a principal altogether?

[Indicator: Years of service, experience in education]

___ years

3B. By the end of this school year, how many years will you have been a principal at this school?

[Indicator: Years of service, experience in education]

___ years

4. What is the highest level of formal education you have completed?

[Indicator: Highest level of education]

- ☐ Did not complete <Upper secondary education—ISCED Level 3>
- ☐ <Upper secondary education—ISCED Level 3>
- ☐ <Post-secondary, non-tertiary education—ISCED Level 4>
- ☐ <Short-cycle tertiary education—ISCED Level 5>
- ☐ <Bachelor's or equivalent level—ISCED Level 6>
- ☐ <Master's or equivalent level—ISCED Level 7>
- ☐ <Doctoral or equivalent level—ISCED Level 8>

5. Have you completed any formal professional training in education administration or management (e.g., leadership or inspection)?

[Indicator: Professional development]

- ☐ Yes
- ☐ No

6A. Which of these options best describes your school?

[Indicator: School ownership (private/public/religious)]

- ☐ Public: managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise.
- ☐ Private/ Independent: managed directly or indirectly by a non-government organization, e.g., a church, trade union, business, or other private institution.

6B. What kind of organization manages your school?

[Indicator: School ownership (private/public/religious)]

- ☐ A church or other religious organization
- ☐ A not-for-profit organization
- ☐ A for-profit organization
- ☐ The government
- ☐ Other

7. Which of the following levels of education does your school provide?

[Indicator: School type/structure (primary/ secondary/language)]

- ☐ <Early childhood education—ISCED Level 0>
- ☐ <Primary education—ISCED Level 1>
- ☐ <Lower secondary education—ISCED Level 2>
- ☐ <Upper secondary education—ISCED Level 3>

8A. How many people live in the city, town, or area where your school is located?

[Indicator: School location (urban/rural)]

- ☐ More than 500,000 people
- ☐ 100,001–500,000 people
- ☐ 50,001–100,000 people

- ☐ 30,001–50,000 people
- ☐ 15,001–30,000 people
- ☐ 3,001–15,000 people
- ☐ 3,000 people or fewer

8B. Which best describes the immediate area in which your school is located?

[Indicator: School location (urban/rural)]

- ☐ Urban—Densely populated
- ☐ Suburban—On fringe or outskirts of urban area
- ☐ Medium size city or large town
- ☐ Small town or village
- ☐ Remote rural

9. Approximately what percentage of students in your school have the following backgrounds?

[Indicator: SES Backgrounds of students]

	0–10%	11–25%	26–50%	51–75%	More than 75%
a) Come from economically disadvantaged homes	☐	☐	☐	☐	☐
b) Come from economically affluent homes	☐	☐	☐	☐	☐

10. What is the total enrollment of students in your school as of <first day of month testing begins, 202x>? (write “0” if there are none)

[Indicator: School size (total enrollment of boys and girls and at target grade)]

___ boys
___ girls

11. What is the total enrollment of <target grade> students in your school as of <first day of month testing begins, 202x>? (write “0” if there are none)

[Indicator: Enrollment & number of classes in target grade]

___ boys
___ girls

12. Approximately what percentage of students in your school have <language of test> as their first language at home?

[Indicator: Language of test/ language of instruction]

- ☐ More than 90%
- ☐ 76–90%
- ☐ 51–75%
- ☐ 26–50%
- ☐ 25% or less

13. How many teachers are at your school? (A full-time teacher is employed at least 90% of the time as a teacher for the full school year.)

[Indicator: Number of teachers/ Staff]

___ full time teachers (> 90% of full-time hours)
___ part time teachers (71-90% of full-time hours)

- ___ part time teachers (50-70% of full-time hours)
- ___ part time teachers (less than 50% of full-time hours)

14. How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following?

[Indicator: School Educational Resources except books (and affection by shortage)]

	Not at all	A little	Some	A lot
a) Educational resource materials (e.g., papers, pencils, materials)	☐	☐	☐	☐
b) Resources for students with learning disabilities	☐	☐	☐	☐

15. How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following?

[Indicator: School human resources (and affected by shortage)]

	Not at all	A little	Some	A lot
a) Qualified teachers	☐	☐	☐	☐

16. Approximately, how many books does your school library have?

[Indicator: School Library (equipment and use)]

- ☐ My school does not have a school library
- ☐ 50 books or fewer
- ☐ 51 to 100 books
- ☐ 101 to 200 books
- ☐ 201 to 500 books
- ☐ 501 to 1000 books
- ☐ 1001 books or more

17. Does your school have the following resources?

[Indicator: Internet]

	Yes	No
a) Reliable internet connection available for teaching and learning	☐	☐

18. Does your school have the following resources?

[Indicator: School facilities (incl. Toilets SDG)]

	Yes	No
a) Electricity	☐	☐
b) Safe drinking water available to students	☐	☐
c) Handwashing facilities with soap and water	☐	☐
d) Facilities accessible to students with disabilities (e.g., ramps, adapted toilets)	☐	☐
e) Availability of toilets or latrines that are separate for girls and boys	☐	☐
f) Sufficient digital devices (computers, laptops, etc.) for school administrative purposes	☐	☐
g) Sufficient digital devices (computers, laptops, etc.) for teaching	☐	☐

19. Does your school have the following classroom resources?*[Indicator: Classroom facilities]*

	Yes	No
a) Sufficient chairs or benches	☐	☐
b) Sufficient desks or tables	☐	☐
c) Cupboards	☐	☐
d) Bookshelves	☐	☐
e) Chalkboard or a white board	☐	☐
f) Smart board	☐	☐
g) Digital projector	☐	☐

20A. How does your school make textbooks available to students in <target grade>?*[Indicator: Availability of textbooks]*

	There are no textbooks for students	One textbook is shared by more than two students	One textbook is shared by two students	Most students have their own textbook, but some have to share	Each student has their own textbook
a) Mathematics	☐	☐	☐	☐	☐
b) <language of test>	☐	☐	☐	☐	☐

20B. How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following?*[Indicator: Availability of textbooks]*

	Not at all	A little	Some	A lot
a) Instructional materials (e.g., textbooks)	☐	☐	☐	☐

21. Does your school provide extra classes or tutoring for students who need additional support in learning?*[Indicator: Provision of additional support in reading or math]*

	Yes, at no additional cost to families	Yes, for a fee	No
a) In reading	☐	☐	☐
b) In mathematics	☐	☐	☐

22. How many digital devices does your school have for use by <target grade> students?*[Indicator: Digital learning resources]*

___ devices

23A. For the <target grade> students in your school: How many days per year is your school scheduled to be open for instruction?*[Indicator: School Year Structure (beginning/end of school year, operational days)]*

___ days

23B. In one calendar week, how many days is the school open for instruction?

[Indicator: School Year Structure (beginning/end of school year, operational days)]

- ☐ 6 days
☐ 5 1/2 days
☐ 5 days
☐ 4 1/2 days
☐ 4 days
☐ Other

24. What is the total instructional time, excluding breaks, on a typical day?

[Indicator: School Year Structure (instructional time per day)]

___ minutes

25. Does your school have double shifts or double sessions? (one group of students attends in the morning and another in the afternoon)

[Indicator: Shifts]

- ☐ Yes
☐ No

26. Within the last 12 month, about how many days was your school closed due to unforeseen circumstances (e.g., strikes, bad weather, etc.)?

[Indicator: Days school closed (Covid, strikes, etc.)]

☐ ___ days

27. How would you characterize each of the following within your school?

[Indicator: Teacher job satisfaction]

	Very high	High	Medium	Low	Very low
a) Teachers' job satisfaction	☐	☐	☐	☐	☐

28. How would you characterize each of the following within your school?

[Indicator: EAS_Instructional and Curricular Leadership and Staff Cohesion]

	Very high	High	Medium	Low	Very low
a) Teacher's understanding of the school's curricular goals	☐	☐	☐	☐	☐
b) Teacher's degree of success in implementing the school's curriculum	☐	☐	☐	☐	☐
c) Teacher's expectations for school achievement	☐	☐	☐	☐	☐

29. How would you characterize each of the following within your school?

[Indicator: EAS_Parental Support]

	Very high	High	Medium	Low	Very low
a) Parental support for student achievement	☐	☐	☐	☐	☐
b) Parental involvement in school activities	☐	☐	☐	☐	☐

30. How would you characterize each of the following within your school?*[Indicator: EAS_Student Engagement and Peer Culture]*

	Very high	High	Medium	Low	Very low
a) Students' regard for school property	☐	☐	☐	☐	☐
b) Students' desire to do well in school	☐	☐	☐	☐	☐

31. To what degree is each of the following a problem among <target grade> students in your school?*[Indicator: SOS_Personal Safety / BULLYING (incl. Teachers)]*

	Not a problem	Minor problem	Moderate problem	Serious problem
a) Intimidation or verbal abuse among students (including texting, emailing, etc.)	☐	☐	☐	☐
b) Physical fights among students	☐	☐	☐	☐
c) Intimidation or verbal abuse of teachers or staff, including online harassment (e.g., messaging, social media, or sharing images or videos online)	☐	☐	☐	☐

32. To what degree is each of the following a problem among <target grade> students in your school?*[Indicator: SOS_School rules and enforcement]*

	Not a problem	Minor problem	Moderate problem	Serious problem
a) Arriving late at school	☐	☐	☐	☐
b) Absenteeism (i.e., unjustified absences)	☐	☐	☐	☐

33. To what degree is each of the following a problem among <target grade> students in your school?*[Indicator: SOS_Student Behavior & Respect]*

	Not a problem	Minor problem	Moderate problem	Serious problem
a) Classroom disturbance	☐	☐	☐	☐
b) Cheating	☐	☐	☐	☐
c) Profanity	☐	☐	☐	☐
d) Vandalism	☐	☐	☐	☐
e) Theft	☐	☐	☐	☐

34. What are your school's policies regarding <grade repetition>?*[Indicator: Policies on grade retention/ academic support & digital devices]*

	Yes	No
a) Students may <repeat a grade> voluntarily (e.g., at the request of or with permission from parents)	☐	☐
b) Students must <repeat a grade> if they do not meet minimum achievement standards at the end of the school year	☐	☐
c) Students may <repeat individual courses> without repeating the entire <grade>	☐	☐

d) Students are expected to leave the school after <repeating grades> a certain number of times	☐	☐
e) <State or district> regulations prohibit <grade repetition>	☐	☐
f) School policies prohibit <grade repetition>	☐	☐

35. To what degree is each of the following a problem among teachers in your school?

[Indicator: Teacher Absenteeism, Teacher availability, Issues at school]

	Not a problem	Minor problem	Moderate problem	Serious problem
a) Arriving late or leaving early	☐	☐	☐	☐
b) Absenteeism	☐	☐	☐	☐
c) Serious behavioral misconduct by teachers or staff (e.g., substance use, inappropriate behavior towards students or colleagues)	☐	☐	☐	☐

36. Does your school provide free meals for students?

[Indicator: Nutrition, meals provided by school]

	Yes, for all students	Yes, for some students	No
a) Breakfast	☐	☐	☐
b) Lunch	☐	☐	☐

8.4 Appendix A4: Full Indicator List

Table 8-1 (Student Level), Table 8-2 (Teacher Level), and Table 8-3 (School Level) list all Indicators identified in the analyzed international background questionnaire versions of the included assessments. They also list the corresponding framework domain classification and the availability of questions loading on the respective indicator. Indicators indicated in blue in the subsequent tables were retained for further analysis.

Table 8-1: Full Indicator List, Student Level

Domain	# ²⁵	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Student Background > Student personal characteristics	1	Gender	✓	X	X	X	X	X	X	X	X		8
Student Background > Student personal characteristics	2	Age	✓	X	X	X	X	X	X	X	X		8
Student Background > Student home environment	3	Language experience	✓	X	X	X	X	X	X	X	X		8

²⁵ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

Domain	#25	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Student Background > Student personal characteristics	4	Immigration Status		X	X					X	X	X	5
Student Background > Student home environment	5	Parents' literacy		(X)		X			X	X	X		5
Student Background > Student home background	6	Parents' Highest formal education		(X)	X		X	X	X	X	X		7
Student Background > Student preparation for learning	7	Participation in formal pre-primary education		(X)	X	X	X	X	X	X	X		8
Student Background > Student personal characteristics	8	Grade repetition			X	X	X	X		X	X		6
Student Content > Student attendance at school	9	Absenteeism	✓	X	X		X			X			4
Student Background > Student home environment	10	Household composition (number of people living at home)	✓		X	X	X	X	X	X		X	7
Student Background > Home learning resources	11	Number of books in the home	✓	X	X	X	X	X	X	X	X		8
Student Background > Home learning resources	12	Household possessions and services	✓	X	(X)	X	X	X	X	X	X	X	9
Student Background > Student home environment	13	Type, material of house			(X)	X	X	(X)		X	X		6
Student Background > Home learning environment	14	Parental involvement in learning (incl. help with schoolwork/ Tuition)			X	X	X	X	X	X	X		7
Student Contexts > Student participation	15	Assignment of Homework			X	X	X	X					4
Student Contexts > Student attitudes/disposition	16	Enjoyment of reading/ mathematics	✓	X		X	X						3

Domain	#25	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
towards specific subjects													
Student Contexts > Student self-confidence	17	Confidence in reading/ mathematics	✓	X	X								2
Student Background > Student preparation for learning	18	Nutrition, meals	✓	X		X	X	X	X				5
Student Background > Student attributes	19	Physical and mental health (incl. Anxiety, Depression)	✓	X		X			X	X		X	5
Student Content > Student's perception of school	20	Well being at school	✓	X	X	X	X	X		X			6
Student Content > Student's perception of school	21	Classroom climate		X	X			X		X			4
Student Content > Student's perception of school	22	Experience of bullying/ harassment		X	X					X	X	X	5
Student Background > Home learning environment	23	Access to books from community/school/class library		X			X						2
Student Contexts > Student participation	24	House chores/ Work			X	X	X	X		X			5
Student Background > Student personal characteristics	25	Commuting to school	✓				X	X		X			3
Student Background > Student preparation for learning	26	Language of class instruction					X			X			2
Student Content > Student's perception of school	27	Perception of Instruction	✓	X	X	X		X		X		X	6
Student -> Student home background, Characteristics and attributes -> Student personal characteristics	28	Student's Grade								X			1
Student -> Student home background, Characteristics and	29	Ethnicity			X								1

Domain	#25	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
attributes -> Student personal characteristics													
Student -> Attained (experiential and learned) -> Student's perception of school	30	Student belonging and identity						X					1
Student -> Student home background, Characteristics and attributes -> Student home environment	31	Parent's Occupation and income		(X)	(X)			(X)		X			4
Student -> Student home background, Characteristics and attributes -> Student preparation for learning	32	Early Learning Activities and Foundational Skills		(X)				(X)					2
Student -> Student home background, Characteristics and attributes -> Student preparation for learning	33	Age of entry into primary education		(X)				(X)					3
Student -> Student (classroom/learning) behaviours -> Student attitudes towards learning	34	Educational aspiration		(X)	X			(X)					3
Student -> Student home background, Characteristics and attributes -> Home learning Resources	35	Student use of digital tools at home/ school		X			X						3
Student -> Student (classroom/learning) behaviours -> Student participation	36	Extra lessons					X						1
Student -> Student home background, Characteristics and attributes -> Home Learning Environment	37	EAS_Parental Support			(X)								1
Student -> Attained (experiential and learned) -> Student's perception of school	38	Parent Perceptions of the school		(X)									1

Domain	#25	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Student -> Student home background, Characteristics and attributes -> Home Learning Environment	39	Parent beliefs			(X)								1
Student -> Student (classroom/learning) behaviours -> Student self-confidence	40	Confidence in ICTS		X									2
Student -> Student (classroom/learning) behaviours -> Student attitudes towards learning	41	Sustainability and Citizenship - Attitudes		X				X					2
Student -> Student (classroom/learning) behaviours -> Student participation	42	Sustainability and Citizenship - Actions		X				X					2
Student -> Student home background, Characteristics and attributes -> Student home environment	43	Sanitation (Washing,...)							X				1
Student -> Attained (experiential and learned) -> Student achievement	44	Resource and knowledge regarding disease (HIV)					X						1
Student -> Attained (experiential and learned) -> Student attendance at school	45	Covid Pandemic		(X)									1
Student -> Student home background, Characteristics and attributes -> Student home environment	46	Neighborhood Safety and Cohesion			(X)								1
Student -> Attained (experiential and learned) -> Student's perception of school	47	Class facilities					X						1
Student -> Attained (experiential and learned) -> Student's perception of school	48	School educational resources			X	X	X						3
Student -> Student (classroom/learning)	49	Reading/ math Activities outside school		X		X	X						3

Domain	# ²⁵	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
behaviours -> Student participation													
Student -> Student (classroom/learning) behaviours -> Student participation	50	General activities outside school			(X)								1

X = indicator available in student questionnaire

(X)=indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

Table 8-2: Full Indicator List, Teacher Level

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Class -> Teacher background and characteristics -> Teacher personal characteristics	1	Teacher Gender		X	X	X	X	X		X	X		7
Class -> Teacher background and characteristics -> Teacher personal characteristics	2	Teacher Age		X	X	X	X	X		X			6
Class -> Teacher background and characteristics -> Teacher's teaching experience	3	Years of Teaching		X	X	X	X	X		X			6
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	4	Highest level of education		X	X	X	X	X		X	X	X	8
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	5	Professional Teaching Qualification			X	X					X		3
Class -> Classroom environment and	6	Subject-Matter Assignments			X		X	X					3

²⁶ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
processes -> Classroom contexts		(general /in sel. class)											
Class -> Classroom environment and processes -> Classroom contexts	7	Grades/ Classes being taught			X	X				X			3
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	8	Initial Teacher Education (duration/quality /Experiences)			X	X	X			X	X		5
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	9	Area(s) of Post-secondary Study / Teacher		X		X		X			X		4
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	10	Professional development (duration, quality)		X	X		X	X		X	X		6
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	11	Confidence in teaching (methods)			X			X					2
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	12	Job satisfaction (incl. General working conditions)		X	X	X	X	X		X			6
Class -> Classroom environment and processes -> Classroom contexts	13	Teacher Absence				X	X			X			3
Class -> Teacher background and characteristics -> Teacher personal characteristics	14	Type of contract with school/ Position			X	X	X			X		X	5
Class -> Teacher background and characteristics -> Teacher personal characteristics	15	Additional paid jobs			X	X				X			3
Class -> Classroom environment and	16	Instructional Resources and				X	X	X		X			4

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
processes -> Classroom environment		Classroom Infrastructure Inventory											
Class -> Classroom environment and processes -> Classroom environment	17	School- or Classroom library and use		X	X		X				X		4
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	18	EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)		X	X	X							3
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	19	EAS_Parental Support		X	X			X		X			4
Class -> Classroom environment and processes -> Classroom environment	20	SOS_Personal Safety (including bullying)		X	X	X							3
Class -> Classroom environment and processes -> Classroom environment	21	SOS_Student Behavior & Respect		X	X			X		X			4
Class -> Classroom environment and processes -> Teachers instructional practice	22	External Supervision & Inspection			X	X	X			X			4
Class -> Classroom environment and processes -> Classroom contexts	23	Class size/ Multi-single grade classroom		X		X				X		X	4
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	24	Teaching conditions/ Teaching Constraints		X	X			X					3
Class -> Classroom environment and processes -> Classroom contexts	25	Language of instruction				X				X		X	3

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Class -> Implemented (operational) -> Teaching and learning	26	Instructional time (Language/ Reading or mathematics)		X		X	X						3
Class -> Implemented (operational) -> Teachers instructional practice	27	Instructional activities (general, Language/ Reading, Math)		X	X			X		X			4
Class -> Implemented (operational) -> Teaching and learning	28	Time spent on activities			X		X						2
Class -> Teacher background and characteristics -> Teacher personal characteristics	29	Native Language of teacher			X					X			2
Class -> Teacher background and characteristics -> Teacher personal characteristics	30	Other occupation of teacher											0
Class -> Teacher background and characteristics -> Teacher personal characteristics	31	Regular payment of teacher				X							1
Class -> Teacher background and characteristics -> Teacher personal characteristics	32	Subjects trained to teach					X						1
Class -> Teacher background and characteristics -> Teachers' qualification and preparation	33	Attitudes towards learning (prejudices)			X								1
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	34	Reading for Enjoyment		X									1
Class -> Teacher background and characteristics -> Teacher personal characteristics	35	Condition of living accommodation					X			X			2

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Class -> Teacher background and characteristics -> Teacher personal characteristics	36	Distance/ Commute time to school					X						1
Class -> Teacher background and characteristics -> Teacher personal characteristics	37	Work in different schools			X					X			2
Class -> Teacher background and characteristics -> Teacher personal characteristics	38	Use of calculators and digital devices		X						X			2
Class -> Implemented (operational) -> Teachers instructional practice	39	Collaboration with external organizations						X					1
Class -> Classroom environment and processes -> Classroom contexts	40	Teacher resource Center					X						1
Class -> Classroom environment and processes -> Classroom environment	41	EAS_Student Engagement and Peer Culture		X				X					2
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	42	SOS_School rules and enforcement		X									1
Class -> Classroom environment and processes -> Classroom environment	43	SOS_School security Policies		X									1
Class -> Classroom environment and processes -> Classroom environment	44	Learning and language difficulties/ sleep/ nutrition		X	X								2

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	45	Highest level of education students will complete			X								1
Class -> Teacher background and characteristics -> Teachers' attitudes and perceptions	46	Parent meetings to discuss student performance and behavior					X						1
Class -> Implemented (operational) -> Teaching and learning	47	Managed to teach the full curriculum				X				X			2
Class -> Implemented (operational) -> Teaching and learning	48	Ability grouping											0
Class -> Implemented (operational) -> Teachers instructional practice	49	Using calculators and digital devices		X									1
Class -> Implemented (operational) -> Teaching and learning	50	Home assignments		X									1
Class -> Implemented (operational) -> Teaching and learning	51	Curriculum coverage		X									1
Class -> Implemented (operational) -> Assessment of learners	52	School assessments		X			X						2
Class -> Implemented (operational) -> Teaching and learning	53	Ask parents to sign homework					X						1
Class -> Implemented (operational) -> Teaching and learning	54	GLOBAL CITIZENSHIP incl. Sustainable development education (SEA-PLM)		X				X					2
Class -> Implemented (operational) -> Teaching and learning	55	HIV					X						1
Class -> Teacher background and characteristics ->	56	Interest, Challenge to fill questionnaire			X								1

Domain	# ²⁶	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
Teachers' attitudes and perceptions													

X - indicator available in student questionnaire

(X) - indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

Table 8-3: Full Indicator List, School Level

Domain	# ²⁷	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics... -> Principals' characteristics	1	Gender			X	X	X	X	X			X	6
School -> School characteristics... -> Principals' characteristics	2	Age			X	X	X	X	X			X	6
School -> School characteristics... -> Principals' characteristics	3	Years of service, experience in education		X	X	X	X	X	X				6
School -> School characteristics... -> Principals' characteristics	4	Highest level of education		X	X	X	X	X	X			X	7
School -> School characteristics... -> Principals' characteristics	5	Professional development			X	X						X	3
School -> School characteristics... -> School characteristics...	6	School ownership (private/public/religious)			X	X	X	X	X	X		X	7
School -> School characteristics... -> School characteristics...	7	School type/structure (primary/secondary/language)			X	X	X			X			4

²⁷ The “#” column gives each indicator’s reference number—a unique identifier used to cross-reference the indicator across tables.

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics... -> Characteristics of surrounding community	8	School location (urban/rural)	✓	X	X	X	X	X	X	X	X		8
School -> School characteristics... -> Student Background	9	SES Backgrounds of students		X	X		X			X	X		5
School -> School characteristics... -> School characteristics...	10	School size (total enrollment of boys and girls and at target grade)	✓	X	X		X	X	X	X	X		7
School -> School characteristics... -> School characteristics...	11	Enrollment & number of classes in target grade		X			X	X	X	X	X		6
School -> School characteristics... -> Student Background	12	Language of test/ language of instruction		X				X	X		X		4
School -> School characteristics... -> School Human resources	13	Number of teachers/ Staff	✓		X	X	X	X	X	X	X		7
School -> School characteristics... -> School Physical resources	14	School Educational Resources except books (and affection by shortage)	✓	X	X							X	3
School -> School characteristics... -> School Human resources	15	School human resources (and affected by shortage)	✓	X	X			X					3
School -> School characteristics... -> School Physical resources	16	School Library (equipment and use)	✓	X	X		X	X	X				5
School -> School characteristics... -> School infrastructure	17	Internet		X	X		X	X	X		X	X	7
School -> School characteristics... -> School infrastructure	18	School facilities (incl. Toilets SDG)	✓	X	X	X	X	X	X	X	X		8

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics... -> School Physical resources	19	Classroom facilities	✓										0
School -> School characteristics... -> School Physical resources	20	Availability of textbooks	✓	X	X			X	X	X			5
School -> School environment and processes -> School support	21	Provision of additional support/ tutoring in reading or math				X		X	X	X			4
School -> School characteristics... -> School Physical resources	22	Digital learning resources	✓	X	X		X	X			X		5
School -> School environment and processes -> School conditions	23	School Year Structure (beginning/end of school year, operational days)	✓	X		X	X					X	4
School -> School environment and processes -> School processes	24	School Year Structure (instructional time per day)	✓	X	X			X					3
School -> School environment and processes -> School conditions	25	Shifts			X	X	X						3
School -> School environment and processes -> School conditions	26	Days school closed (Covid, strikes, etc.)	✓	X		X	X			X		X	5
School -> School environment and processes -> School climate	27	Teacher job satisfaction	✓										0
School -> Implemented (perceived) -> Instructional leadership	28	EAS_Instructional and Curricular Leadership and Staff Cohesion	✓	X	X			X					3

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School environment and processes -> School support	29	EAS_Parental Support	✓	X	X			X				X	4
School -> School environment and processes -> School climate	30	EAS_Student Engagement and Peer Culture	✓	X				X					2
School -> School environment and processes -> School climate	31	SOS_Personal Safety / BULLYING (incl. Teachers)	✓	X	X		X	X		X	X		6
School -> School environment and processes -> School processes	32	SOS_School rules and enforcement	✓	X			X						2
School -> School environment and processes -> School climate	33	SOS_Student Behavior & Respect	✓	X	X		X	X				X	5
School -> School environment and processes -> School processes	34	Policies on grade retention/ academic support & digital devices								X			1
School -> School environment and processes -> School conditions	35	Teacher Absenteeism, Teacher availability, Issues at school	✓	X		X	X			X			4
School -> School environment and processes -> School support	36	Nutrition, meals provided by school	✓				X		X	X			3
School -> School characteristics and surrounding community characteristics -> Principals' characteristics	37	Degree				X							1
School -> School environment and processes -> School conditions	38	Type of contract			X								1

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics and surrounding community characteristics -> Teachers' characteristics	39	Training and teaching degrees		X	X		X						3
School -> School environment and processes -> School support	40	Training needs of teacher			X								1
School -> School characteristics and surrounding community characteristics -> Teachers' characteristics	41	Other occupation			X								1
School -> School environment and processes -> School climate	42	Job satisfaction							X				1
School -> School environment and processes -> School climate	43	Beliefs			X								1
School -> School environment and processes -> School processes	44	Teaching hours					X						1
School -> School characteristics and surrounding community characteristics -> School characteristics and surrounding community characteristics	45	Multi-single grade classroom			X	X							2
School -> School environment and processes -> School processes	46	Ability grouping			X					X			2
School -> School characteristics and surrounding community characteristics -> Student Background	47	Race			X								1

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics and surrounding community characteristics -> Student Background	48	Language of students			X								1
School -> School characteristics and surrounding community characteristics -> Characteristics of surrounding community	49	Facilities school community				X	X	X					3
School -> School characteristics and surrounding community characteristics -> Characteristics of surrounding community	50	Community environment			X					X			2
School -> School environment and processes -> School support	51	Community and Parental support					X			X			2
School -> School environment and processes -> School context	52	Partnerships with community and international associations/org anizations				X				X			2
School -> School characteristics and surrounding community characteristics -> Student Background	53	School Readiness		X						X			2
School -> School environment and processes -> School conditions	54	Accessibility of school (school closure)				X							1
School -> School environment and processes -> School support	55	Provision of additional support in other subjects											0
School -> School environment and processes -> School support	56	Support from local community					X						1

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School characteristics and surrounding community characteristics -> School Physical resources	57	Online learning management system		X									1
School -> School environment and processes -> School processes	58	Arrangements for teacher absences				X							1
School -> School environment and processes -> School leadership	59	School inspection / teacher performance evaluation				X	X	X					3
School -> School characteristics and surrounding community characteristics -> School Physical resources	60	Availability of official curricula				X							1
School -> School environment and processes -> School processes	61	Student grade promotion				X							1
School -> School environment and processes -> School climate	62	Students receive encouragements / awards				X							1
School -> School environment and processes -> School conditions	63	SOS_School security Policies											0
School -> School environment and processes -> School climate	64	Wellbeing (teachers, students)						X					1
School -> Implemented (perceived) -> School educational emphases	65	Emphasis on Environmental Sustainability		X									1
School -> Implemented (perceived) -> School-based curriculum	66	Grade of first emphasis of reading skills		X									1

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
School -> School environment and processes -> School leadership	67	School leadership activities			X	X	X						3
School -> School environment and processes -> School processes	68	School Action Plan				X							1
School -> School environment and processes -> School processes	69	Pedagogical days											0
School -> School characteristics and surrounding community characteristics -> School characteristics and surrounding community characteristics	70	Funding, school fees				X				X			2
School -> School environment and processes -> School processes	71	Effort from school for environmental sustainability		X									1
School -> School characteristics and surrounding community characteristics -> Teachers' characteristics	72	Teacher qualifications/ teacher training					X	X	X				3
School -> School environment and processes -> School conditions	73	Student Absence							X				1
School -> School environment and processes -> School processes	74	Student attendance monitoring						X					1
School -> School environment and processes -> School support	75	Professional development for teachers						X	X				2
School -> School environment and processes -> School processes	76	Parent meetings and informing parents				X							1

Domain	#27	Indicator	In LaNA	TIPI	ERCE	PASEC	SACMEQ	SEA-PLM	AMPL	PISA-D	SDG	WB	Score
--------	-----	-----------	---------	------	------	-------	--------	---------	------	--------	-----	----	-------

X - indicator available in student questionnaire

(X) - indicator only available in parent questionnaire (applicable to TIMSS/PIRLS/ERCE/SEA-PLM)

Note: 3 indicators (AIDS related/ # of students and teachers died, Covid Pandemic, Questionnaire rating) were excluded from this project due to limited cross-comparability and relevance.

8.5 Appendix A5: Additional Indicators suggested by Country Experts

Table 8-4 summarizes input from three LaNA NRCs on additional indicators that are policy relevant in their country contexts and are therefore suggested for inclusion in LMIC-specific background instruments.

Table 8-4: Additional Indicators suggested by Country Experts

Level	Palestine	Burkina Faso	Senegal
Students	The issue of e-books		1. Pedagogical support out-of-class (Reinforcement course) 2. Possession of Individual Textbooks (Q12 lists household goods and Q23 access to the library, but own possession of the math or French textbook is crucial.) 3. Child Labor (Seasonal Details) 4. Use of ICT (Information and Communication Technology) 5. Type of Establishment (Self-Declaration or Code)
Teachers		Pedagogical support for the teacher by the school director or the school team	1. Travel Time and Teacher Accommodation 2. Formative Assessment Practice 3. Membership of a Pedagogical Animation Unit (CAP)
Schools	Availability of educational aids	Is your school fenced? Is there shade in the schoolyard?	1. Management of Special Classes (Double Flow) and Multigrade Classes 2. Local Governance (School Management Committee – CGE) 3. Access to Water and Sanitation (Aligning with the SDGs)

8.6 Appendix A6: Overview on Selected and Modifications

Table 8-5 (Student Level), Table 8-6 (Teacher Level), and Table 8-7 (School Level) list the original questionnaire items identified, along with their location in the source questionnaire, the final suggested item wording, the modification level, and the justification for the adaptation.

Table 8-5: Student Questionnaire Selected Items and Modifications

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Gender	A	Are you a girl or a boy? Girl Boy	Which of these describes you? Girl Boy <Other>	LaNA:StQ01	2	Stem refined for neutrality; addition of "<Other>" category shifts the binary scheme and may affect response distributions.
Age	A	How old are you? ___ years old	How old are you? ___ years old	LaNA:StQ02	1	Identical stem and direct numeric adoption; psychometric properties and scaling parameters remain stable.
Language experience	A	How often do you speak the language of the test at home? Always Almost always Sometimes Never	How often do you speak the <language of the test> at home? Always Almost always Sometimes Never	LaNA:StQ03	1	Maintained standard IEA 4-point frequency scale; high stability for familiarity measure.
Language experience	B	What language to you speak at home most of the time? (OPTIONAL) <language 1> <language 2> <language 3> <other> (Note: If two languages are spoken at the same frequency, choose the one you learnt first.)	What language do you speak at home most of the time? <language of test> <language 2> <language 3> <Other> (Note: If two languages are spoken at the same frequency, choose the one you learned first.)	SEA-PLM:StQ04	1	Identical intent to complement the construct in multilingual settings.
Immigration Status	A	Were you born in <country>? Yes No	Where were you born? In this same village/town/city where I live now In another village/town/city in <this country> In another country I don't know	TIMSS:SQG07	3	Significant construct shift: Transformed binary Yes/No into a 4-point categorical displacement scale; recalibration required.
Immigration Status	B	Was your mother/father born in <country>?		TIMSS:SQG06; ERCE:QA6:Q07,Q08	3	NEW dimension: Addition of parental displacement details following World Bank suggestions.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Parents' literacy	A	Among the people who live in your home, who can read? (mark Yes/No for each) Mother Yes/No; Father Yes/No; Male guardian Yes/No; Female guardian Yes/No; Brothers/Sisters Yes/No; Grandparents/Uncles/Aunts Yes/No; Another person Yes/No	How many adults in your home can read or write? No adults at home can read or write One adult at home can read or write More than one adult at home can read or write I don't know	PASEC:QE6_21	3	Significant shift: Construct moved from identifying specific literate household members to a categorical count of literate adults.
Parents' Highest formal education	A	What is the highest level of education that your father (or male guardian) /mother (or female guardian) have completed? <ISCED Level 6> or higher Yes/No; <ISCED Level 4 or 5> Yes/No; <ISCED Level 3> Yes/No; <ISCED Level 2> Yes/No; <ISCED Level 1> Yes/No; <Religious education> Yes/No; Did not go to school <or attend religious education> Yes/No; He / she doesn't have a mother / female guardian Yes/No;	What is the highest level of education that your father (or male guardian) /mother (or female guardian) have completed? a) Did not go to school b) <Primary education—ISCED Level 1> c) <Lower secondary education—ISCED Level 2> d) <Upper secondary education—ISCED Level 3> e) <Post-secondary non-tertiary education or short-cycle tertiary education—ISCED Level 4 or 5> f) <Bachelor's or equivalent level—ISCED Level 6> g) <Master's or equivalent or doctoral or equivalent level—ISCED Level 7 or 8> g) I don't know h) Not applicable Mother or female caregiver Father or male caregiver	SEA-PLM:StQ11, StQ12	2	Category refinement: Standard ISCED levels kept, but adding "I don't know" and "Not applicable" affects response distributions.
Participation in formal pre-primary education	A	Did you go to a pre-primary or kindergarten before starting primary school? Yes, for 2 years or more	Did you go to a preschool or kindergarten before starting primary school? Yes, for 2 years or more Yes, for between 1 year and 2	SEA-PLM:StQ05	2	Significant expansion: Expanded response options to include more granular duration and uncertainty options.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		Yes, for 1 year No	years Yes, for less than 1 year No I don't know			
Grade repetition	A	Have you ever repeated a grade? Yes No	Have you ever repeated a grade? Yes No	SEA-PLM:StQ13	1	Direct adoption of simple binary grade repetition measure; high stability expected.
Absenteeism	A	About how often are you absent from school? Never or almost never Once a month Once every two weeks Once a week or more	About how often are you absent from school? Never or almost never About once every two months About once per month About once every two weeks About once a week or more	LaNA:StQ08	2	Refinement of frequency brackets for better granularity (added "About once every two months").
Absenteeism	B	During the past week, how many days were you late for school? None 1 day 2 days 3 days 4 or more I don't know/ I don't remember	About how often do you arrive late to school? Never or almost never About once every two months About once per month About once every two weeks About once a week or more	ERCE: QA6:Q18	3	NEW scale: Shifted punctuality from a count of days to a standardized frequency scale.
Household composition	A	How many people live in your home? (Do not count yourself.) 1–3 4–6 7–9 10 or more	How many people currently live in your home? (Including yourself.) 1–3 4–6 7–9 10 or more	LaNA:StQ06	2	Minor adjustment: Clarification to "Including yourself" may shift numeric counts, but the core construct remains similar.
Household composition	B	How many siblings (brothers and sisters) do you have? None One Two Three Four More than four	How many siblings (brothers and sisters) do you have? a) None b) One c) Two d) Three e) Four f) More than four	SEA-PLM:StQ03	3	NEW dimension: Introduced a sibling count to determine dependency ratios.
Number of books in the home	A	About how many books are there in your home? (Do not count	How many books are there in your home? (Do not count	LaNA:StQ04	2	Refinement of ranges: Brackets adjusted to match

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		schoolbooks, newspapers, magazines) 0–10 11–25 26–100 101–200 More than 200	schoolbooks, newspapers, magazines) No books 1–10 11–20 21–30 31–50 51–100 More than 100			ERCE/AMPL lower floor to better fit LMIC contexts.
Number of books in the home	B	How many books are in your home? (Consider all types of books: poetry, novels, dictionaries, study books, etc.) No books There are 10 or less There are between 11 and 20 There are between 21 and 30 There are between 31 and 50 There are between 51 and 100 There are more than 100.				No secondary question provided.
Household possessions and services	A	Do you have any of these things at your home? (Tick Yes/No for each) - Running water - Electricity - Refrigerator - Mobile phone - Computer or tablet Yes No (for each item)	Do you have any of these things at your home? a) Running water b) Electricity c) Refrigerator (fridge) d) Mobile phone e) Computer or tablet f) Access to the internet g) Study desk/table for your use h) Your own <bedroom> i) Television j) Stove for cooking food k) Car or motorcycle l) Flush toilet m) <country-specific> Yes No	LaNA:StQ05	3	Significant expansion: Massive addition of asset items (WASH, transport, etc.) to meet modern SDG and World Bank standards.
Household possessions and services	B	Do you have any of these at home? (Tick Yes/No for each)		TIMSS:SQG05		No secondary question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		- A table/desk for studying - Electricity - Television/radio - Refrigerator - Internet access - Computer or tablet - Mobile phone Yes No				
Type, material of house	A	The walls of your house are made of... Cardboard Grass Mud & sticks Stones Metal sheets Wood Cut stones	What are the walls of your home mostly made of? Cardboard Grass Mud and sticks Stones Metal sheets Wood Cut stones or bricks Other	SACMEQ:PWALL	3	Significant shift: Moved from specific materials to simplified MICS wall quality levels (natural, rudimentary, finished).
Parental involvement in learning	A	Do you receive help to do your homework? Yes No	How often does someone at home help you with your homework? I do not get any homework Never Sometimes Most of the time	PASEC:QE6_23	3	Scale shift: Transformed from binary "Yes/No" to a frequency scale including a category for not receiving homework.
Parental involvement in learning	B	If yes, how often do you get help? Always often sometimes never		PASEC:QE6_25 categories		No secondary question provided.
Assignment of Homework	A	How often are you usually given homework? I do not get any homework Once or twice each month Once or twice each week Most days of the week	About how often do you get homework? a) In mathematics b) In <language of test>/reading I do not get any homework Once or twice a month Once or twice a week Most days of the week	PASEC:QE6_22	2	Category refinement: Standard frequency scale maintained but split into subject-specific tracks (Math and Language).

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Assignment of Homework	B	(DOES TEACHER CHECK HOMEWORK QUESTION) MAY NEED TO BE ADDED				No secondary question provided.
Enjoyment of reading/ mathematics	A	What do you think about reading? - a) I like talking about what I read with other people - b) I would be happy if someone gave me a book as a present - c) I think reading is boring - d) I would like to have more time for reading - e) I enjoy reading - f) I learn a lot from reading - g) I like to read things that make me think - h) I like it when a book helps me imagine other worlds Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree with these statements about reading? a) I like talking about what I read with other people b) I would like to have more time for reading c) I enjoy reading d) I like to read things that make me think Agree a lot Agree a little Disagree a little Disagree a lot	PIRLS:SQR07a-h	2	Item set reduction (Reading): Reduced set for brevity while maintaining the standard 4-point agreement scale.
Enjoyment of reading/ mathematics	B	How much do you agree with these statements about learning mathematics? - a) I enjoy learning mathematics - b) I wish I did not have to study Mathematics - c) Mathematics is boring - d) I learn many interesting things in mathematics - e) like mathematics - f) like any schoolwork that involves numbers - g) I like to solve mathematics problems - h) I look forward to mathematics lessons - i) Mathematics is one of my favorite subjects	How much do you agree with these statements about learning mathematics? a) I enjoy learning mathematics b) I like mathematics c) I look forward to mathematics lessons d) Mathematics is one of my favorite subjects Agree a lot Agree a little Disagree a little Disagree a lot	TIMSS:SQMS02a-i	2	Item set reduction (Math): Parallel reduction for brevity; core constructs remain same but scale properties may shift.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		Agree a lot Agree a little Disagree a little Disagree a lot				
Confidence in reading/ mathematics	A	How well do you read? Tell how much you agree with each of these statements - a) I usually do well in reading - b) Reading is easy for me -c) I have trouble reading stories with difficult words - d)Reading is harder for me than for many of my classmates -e)Reading is harder for me than any other subjects -f) I am just not good at reading Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree with these statements about reading? a) I usually do well in reading b) Reading is easy for me c) I understand most of what I read Agree a lot Agree a little Disagree a little Disagree a lot	PIRLS:SQR08	2	Item set selection (Reading): Selected items for brevity while maintaining standard 4-point agreement.
Confidence in reading/ mathematics	B	How much do you agree with these statements about mathematics? - a) I usually do well in mathematics b) Mathematics is harder for me than for many of my classmates c)I am just not good at mathematics - d) Mathematics is easy for me - e) I am good at working out difficult mathematics problems f) Mathematics is harder for me than any other subject g) Mathematics makes me confused Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree with these statements about mathematics? a) I usually do well in mathematics b) Mathematics is easy for me c) I am good at working out difficult mathematics problems Agree a lot Agree a little Disagree a little Disagree a lot	TIMSS:SQMS05	2	Item set selection (Math): Parallel selection for brevity; core constructs remain same but scale properties may shift.
Nutrition, meals	A	On a normal school day, which of the following meals do you have? - breakfast	Which meals do you eat on a normal school day? a) Breakfast	SEA-PLM:StQ07	3	NEW dimension: Added a categorical meal checklist for food security profiling.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		- Lunch - Dinner Yes No	b) Lunch c) Dinner Yes No			
Nutrition, meals	B	How often do you feel this way when you arrive at school? I feel hungry Every day Almost every day Sometimes Never	How often do you have trouble paying attention in school because of these things? a) I feel hungry Every day Almost every day Sometimes Never	LaNA:StQ09b	3	Significant construct shift: Shift from health status (feeling hungry) to its impact on classroom attention.
Physical and mental health	A	How often do you feel tired when you arrive at school? (LaNA) Every day Almost every day Sometimes Never	How often do you have trouble paying attention in school because of these things? a) I feel tired Every day Almost every day Sometimes Never	LaNA: ST09a	3	Significant shift: Construct moved from physical status to “trouble paying attention” due to tiredness.
Physical and mental health	B	The next five statements are about your health I can see what is written on the board without difficulty I can hear the teacher’s voice clearly when he or she is giving a lesson I have a physical disability that makes it difficult for me to walk or use stairs. I have a physical disability that makes it difficult for me to grasp small objects like a pencil or scissors. I often get so sick I cannot play, work or go to school. Yes No	Do you agree with these statements about your health: a) I can clearly see the words the teacher writes on the board b) I can clearly hear the teacher when he or she is speaking c) I have a physical impairment that makes it hard for me to walk or go up and down stairs d) I have a physical impairment that makes it hard for me to hold or use small things like a pencil or scissors e) I get sick so often that sometimes I cannot play, do things, or go to school	PISA_D:StQ07	3	NEW dimension: Added a detailed health and disability impairment agreement scale.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			Yes No			
Well being at school	A	How much do you agree with these statements? - I like being in school - I feel safe when I am at school - I feel like I belong at this school Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree with these statements about your school? a) I like being in school b) I feel safe when I am at school c) I feel like I belong at this school Agree a lot Agree a little Disagree a little Disagree a lot	LaNA:StQ10	1	Identical wording: Core items (Like, Safe, Belong) maintained with standard 4-point agreement; high stability.
Classroom climate	A	How often do these things happen in your reading lesson? - Students don't listen to what the teacher says - There is too much noise for students to work well - My teacher has to wait a long time for students to be quiet - Students interrupt the teacher - My teacher has to keep telling us to follow the classroom rules Often Sometimes Rarely Never	How often do these things happen in your class? a) Students don't listen to what the teacher says b) There is too much noise for students to work well c) My teacher has to wait a long time for students to be quiet d) Students interrupt the teacher e) My teacher has to keep telling us to follow the classroom rules Every or almost every day Once or twice a week Once or twice a month Never or almost never	PIRLS:SQR02	2	Stem adjustment: Generalized from subject-specific (Reading) to the general classroom context.
Experience of bullying/harassment	A	COPY PIRLS IN: -> SUGGESTION TO CHANGE TO THESES OPTIONS AND LAST MONTH Never Sometimes Many times Every day or almost every day	During this year, how often have other students from your school done any of the following things to you (in person, through messaging, or through social media)? a) Made fun of me or called me names	PIRLS:SQG11	1	Direct adoption of the full standard 10-item PIRLS bullying scale; parameters remain stable.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			b) Left me out of their games or activities c) Spread lies about me d) Stole something from me e) Damaged something of mine on purpose f) Hit or hurt me g) Made me do things I didn't want to do h) Sent me nasty or hurtful messages online i) Shared nasty or hurtful information about me online j) Threatened me At least once a week Once or twice a month A few times Never			
Access to books from community/school/class library	A	How often do you borrow books (including ebooks) from your <school or local library>? At least once a week Once or twice a month A few times a year Never or almost never	How often do you borrow books (including ebooks) from your <class/ school or local library>? I do not have access to a <class/ school or local library> At least once a week Once or twice a month A few times a year Never or almost never	PIRLS:SQR04	1	Direct adoption of the standard PIRLS borrowing frequency stem.
House chores/ Work	A	How often do you do the following activities when you are not at school? House chores/Farm work/Commercial activities/Physical work/ combine e+f: Taking care of family members Never or hardly never monthly weekly daily or almost daily	How often do you do these things outside of school? a) Do house chores (e.g., cleaning, cooking, fetching water or firewood) b) Do farm work (e.g., planting, harvesting, caring for animals) c) Help to sell things or earn money (e.g., selling at a market, helping in a shop) d) Take care of family	SEA-PLM:StQ06	2	Standardization: Response options moved to a 4-point frequency scale to match universal IEA styles.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			members (e.g., looking after siblings or sick relatives) Every day Almost every day Sometimes Never			
House chores/ Work	B	Do you do farm or paid work when not at school?		PASEC:QE6_17–19		No secondary question provided.
Commuting to school	A	How long does it take you to get to school each day? 30 minutes or less 30 minutes up to 1 hour 1 hours to 2 hours 2 hours or more	How long does it take you to get to school each day? Less than 30 minutes 30 minutes up to 1 hour 1 hour to 2 hours 2 hours or more	LaNA:StQ07	1	Identical categories: Standard 4-point duration brackets maintained from original LaNA source.
Language of class instruction	A	How often does your teacher use a language different from the test language? Never Sometimes Often	How often does your reading/ language teacher use a language different from the <language of test> when teaching? Never Sometimes Often	SACMEQ:PTLANG	3	Significant shift: New structural requirement to identify specific secondary languages used in class.
Language of class instruction	B	What language did most of your teachers use for instruction in <ISCED 1>? <Language 1> <Language 2> <Language 3> < ...etc. > Other language	Which other language does your language teacher use the most when teaching? None, my teacher only uses the <language of test> <language 1> <language 2> <language 3> <Other>	PISA-D:StQ12	3	NEW dimension: Mirror of home language checklist to identify classroom linguistic exposure.
Perception of Instruction	A	How much do you agree with these statements about your reading lessons? - I know what my teacher expects me to do. - My teacher is easy to understand. - I am interested in what my teacher says. Agree a lot	How much do you agree with these statements about your reading lessons? a) I know what my teacher expects me to learn in each lesson b) My teacher is easy to understand c) I am interested in what my	LaNA:StQ13	3	Significant expansion (Reading): Structural increase from 3 to 5 items to better capture teacher practices.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		Agree a little Disagree a little Disagree a lot	teacher says d) My teacher asks me to explain my answers e) My teacher has clear answers to my questions Agree a lot Agree a little Disagree a little Disagree a lot			
Perception of Instruction	B	How much do you agree with these statements about your mathematics lessons? - I know what my teacher expects me to do. - My teacher is easy to understand. - I am interested in what my teacher says. ☐ Agree a lot ☐ Agree a little ☐ Disagree a little ☐ Disagree a lot	How much do you agree with these statements about your mathematics lessons? a) I know what my teacher expects me to learn in each session b) My teacher is easy to understand c) I am interested in what my teacher says d) My teacher asks me to explain my answers e) My teacher is good at explaining mathematics Agree a lot Agree a little Disagree a little Disagree a lot	LaNA:StQ14	3	Category refinement (Math): Increase in response options while the stem remains unchanged; thresholds may shift.

Table 8-6: Teacher Questionnaire Selected Items and Modifications

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Teacher Gender	A	Which of these describes you? Female Male	Which of these describes you? Female Male <Other>	PIRLS:TQG02; TIMSS:TQG02	2	Addition of category for inclusivity alters the binary response scheme; thresholds may shift.
Teacher Age	A	How old are you? Under 25. 25–29. 30–39. 40–49. 50–59. 60 or more.	How old are you? Under 25 25–29 30–39 40–49 50–59 60 or more	PIRLS:TQG03; TIMSS:TQG03	1	Identical stem and standardized age bands from TIMSS; properties and parameters remain stable.
Years of Teaching	A	By the end of this school year, how many years will you have been teaching altogether? _____ years (Please round to the nearest whole number.)	By the end of this school year, how many years will you have been teaching altogether? _____ years (Please round to the nearest whole number.)	PIRLS:TQG01; TIMSS:TQG01	1	Identical stem and numeric input for career total; properties remain stable.
Years of Teaching	B	How many years of work experience do you have? Year(s) working as a teacher at this school. ____years	By the end of this school year, how many years will you have been teaching at this school? _____ years (Please round to the nearest whole number.)	PISA_D: TQ18B	3	NEW dimension: Added secondary question to track school-specific tenure alongside total career years.
Highest level of education	A	What is the highest level of formal education you have completed? Did not complete <Upper secondary education—ISCED Level 3> <Upper secondary education—ISCED Level 3> <Post-secondary, non-tertiary education—ISCED Level 4> <Short-cycle tertiary education—ISCED Level 5>	What is the highest level of formal education you have completed? Did not complete <Upper secondary education—ISCED Level 3> <Upper secondary education—ISCED Level 3> <Post-secondary non-tertiary education—ISCED Level 4> <Short-cycle tertiary education—ISCED Level 5> <Bachelor's or equivalent	PIRLS:TQG04; TIMSS:TQG04	1	Maintained standard international ISCED classifications; scaling parameters from original studies can be used directly.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		<Bachelor's or equivalent level—ISCED Level 6> <Master's or equivalent level—ISCED Level 7> <Doctor or equivalent level—ISCED Level 8>	Level—ISCED Level 6> <Master's or equivalent level—ISCED Level 7> <Doctoral or equivalent level—ISCED Level 8>			
Professional Teaching Qualification	A	Do you have a school teaching degree? Yes. No.	Do you have a formal teaching qualification that qualifies you to teach in schools? (e.g., a teacher training certificate, diploma, or degree, completed either as part of or in addition to your highest level of education) Yes No	ERCE:PQ6:Q14	2	Stem refined to explicitly distinguish a “formal qualification” from a general degree for higher precision.
Subject-Matter Assignments (general /in sel. class)	A	Regarding students in the <classroom assessed>, what subjects do you teach? Mathematics. Science. Language. (Mark all that apply.)	Which of the following two subjects do you teach in the selected class? a) Mathematics b) <language of test> Yes No	ERCE:PQ6:Q07	2	Refined focus: Shifted from a general subject checklist to focus specifically on the two core assessed subjects (Math/Language).
Grades/ Classes being taught	A	Regarding students in the <classroom assessed>, what grade(s) do you teach? Only in third grade. Only in sixth grade. In the third and sixth years.	Which <grade(s) or class level(s)> do you currently teach in this school? <Grade or class level 1> <Grade or class level 2> <Grade or class level 3> <Grade or class level 4> <Grade or class level 5> <Grade or class level 6> <Other>	ERCE:PQ6:Q06	3	Significant shift: Expanded scope from identifying the specific assessed classroom to all grades taught in the school.
Grades/ Classes being taught	B	What <grades> do you teach in this school?		PISA_D:TQ04		Question 2 merged into Suggestion 1.
Initial Teacher Education (duration/quality/Experiences)	A	How many years did you study to become a school teacher? _____ year(s).	Did you complete any pre-service teacher training? (This may have been completed as part of a regular degree or as	ERCE:PQ6:Q16	3	Significant scale change: Shift from numeric “years” of study to a categorical duration

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			separate teacher training). No Yes, for up to 1 year Yes, for more than 1 and up to 2 years Yes, for more than 2 and up to 3 years Yes, for more than 3 years			scale including a “No training” option.
Initial Teacher Education (duration/quality/Experiences)	B	During your <post-secondary> education, what was your major or main area(s) of study? 8.a) Education - Primary/Elementary 8.b) Education-Secondary 8.c) Mathematics 8.d) Science 8.d) <language of test> 8.e) other Yes. No.		TIMSS:TQG5 A		Question 2 merged into Suggestion 1.
Area(s) of Post-secondary Study / Teacher	A	During your <post-secondary> education, what was your major or main area(s) of study? 9.a) Education - Primary/Elementary 9.b) Education-Secondary 9.c) Mathematics 9.d) Science 9.e) <language of test> 9.f) other Yes. No.	During your <post-secondary> education, what was your major or main area(s) of study? (If you did not complete any <post-secondary> education, mark “No <post-secondary> education”). a) No <post-secondary> education b) Education—Primary/Elementary c) Education—Secondary d) Mathematics e) <language of test> f) Other Yes No	TIMSS:TQG5A	2	Added “No post-secondary education” category, which affects potential response distributions and calibration.
Area(s) of Post-secondary Study / Teacher	B	If your major or main area of study was education, did	If your major or main area of study was education, did you	TIMSS:TQG5 B	1	Negligible change; maintained the standard

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		you have a <specialization> in any of the following? 9.a) Mathematics 9.b) Science 9.c) Language 9.d) Other Yes. No.	have a specialization in any of the following? a) Mathematics b) <language of test> c) Other subject Yes No			specialization checklist from PIRLS/TIMSS.
Professional development (duration, quality)	A	In the past two years, have you participated in professional development in any of the following? (List includes content, pedagogy, technology, assessment, etc.) Yes. No.	In the past two years, have you participated in formal professional development in the following areas, and how important are these areas for your current teaching position? a) Teaching reading comprehension skills or strategies b) Integrating reading literacy across the curriculum c) Teaching mathematics content and problem-solving d) Pedagogical methods for teaching reading or mathematics e) Integrating technology into instruction f) Assessing students' learning in reading or mathematics g) Addressing students' diverse language, learning, or inclusion needs h) Differentiating instruction based on student levels and interests Participated: Yes No Importance: Not important Important Very important	TIMSS:TQM09A/TQS12A; PIRLS:TQG07A	3	Significant dimension shift: Added an entirely new "Importance" rating dimension to the standard participation checklist.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Professional development (duration, quality)	B	If you answered Yes to some of the activities, how much did your <professional improvement> as a teacher contribute to your professional development? 5-point scale: It did not contribute anything... to It contributed a lot to my performance as a teacher.	If you answered yes to any of the activities above, to what extent did these formal professional development activities improve your teaching practices? A lot Some A little Not at all	ERCE:PQ6:Q22	2	Scale standardization: Quality categories adjusted to match standard 4-point frequency scales.
Confidence in teaching (methods)	A	In your teaching, how confident or secure do you feel in performing the following activities? 11.a) To have my students develop critical thinking skills (e.g., debate, evaluate different opinions and substantiate their own). 11.b) I pose interesting or challenging questions that make my students think 11.c) Enhance students' social-emotional skills (e.g., empathy, self-management, tolerance). 11.d) Explain difficult content in several different ways (e.g., using a variety of examples, metaphors, diagrams, models, or instructional materials). 11.e) Use varied teaching strategies and alternate among them (e.g., discussions, projects, dramatization, problem	How confident are you in doing each of these things when teaching this class? a) Having my students develop critical thinking skills (e.g., debate, evaluate different opinions and substantiate their own) b) Posing interesting or challenging questions that make my students think c) Enhancing students' social-emotional skills (e.g., empathy, self-management, tolerance) d) Explaining difficult content in several different ways (e.g., using a variety of examples, metaphors, diagrams, models, or instructional materials) e) Using varied teaching strategies and alternating among them (e.g., discussions, projects, dramatization, problem solving, etc.) Very high High Medium Low	ERCE:PQ6:Q39	2	Response category adjustment: Categories changed from ERCE "Quite/Very" to standard TIMSS "High/Medium".

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		solving, etc.). Little or nothing Something Quite Very				
Job satisfaction (incl. General working conditions)	A	How often do you have these feelings about being a teacher? 12.a) I am content with my profession as a teacher 12.b) I find my work full of meaning and purpose 12.c) I am enthusiastic about my job 12.d) My work inspires me 12.e) I am proud of the work I do 12.f) I feel appreciated as a teacher 12.g) I enjoy the challenges of teaching Very often Often Sometimes Never	How often do you have these feelings about being a teacher? a) I am content with my profession as a teacher b) I find my work full of meaning and purpose c) I am enthusiastic about my job d) My work inspires me e) I am proud of the work I do f) I feel appreciated as a teacher g) I enjoy the challenges of teaching Very often Often Sometimes Never	TIMSS:TQG08	1	Direct adoption of the standard 7-item TIMSS job satisfaction scale; parameters remain stable.
Teacher Absence	A	During the last two months, how many days have you been absent apart from holidays? _____ days	During the last month, did any of the following situations prevent you from going to work? a) Health-related reasons (e.g., illness, mental or emotional health problems, medical appointments) b) Family-related reasons (e.g., caring for family members, family emergency) c) Workplace or labor-related reasons (e.g., conflicts or violence at school, strike) d) Transport. weather, or other	PASEC:QM_14	3	Significant format change: Shift from a numeric day count of absences to a categorical checklist of specific reasons.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			external reasons (e.g., lack of transport, extreme weather or hazards) Yes, No			
Teacher Absence	B	During the last month, did any of the following situations prevent you from going to work? yes no		PISA: TQ33		No secondary question provided.
Type of contract with school/ Position	A	What type of employment relationship (contract) do you have with this school? For an indefinite term (tenured or appointed position) For a term of one year or more with the possibility of continuing (interim, contract) For a term of less than one year (substitutes, etc.) Other	How long is your employment contract with this school? For an indefinite term (tenured or appointed position) For a term of one year or more with the possibility of continuing (interim, contract) For a term of less than one year (substitutes, etc.) Other	ERCEPQ6:Q09	3	Shift in focus: Although options are similar to ERCE, the construct shifted from type of contract to duration focus.
Additional paid jobs	A	In addition to teaching in school, how many hours per week do you work in another job that is not related to teaching? None. Up to 10 hours per week. Between 10 and 19 hours per week. Between 20 and 30 hours per week. More than 30 hours per week.	In addition to your regular teaching at this school, how many hours per week do you spend on other paid work (including private tutoring or any other job)? None Fewer than 10 hours per week 10 to 19 hours per week 20 to 30 hours per week More than 30 hours per week	PISA_D:TQ16	2	Stem clarified to include private tutoring; range labels standardized for improved cross-country comparison.
Instructional Resources and Classroom Infrastructure Inventory	A	Which of the following items do you have in your classroom?	Which of the following items do you have in your classroom? a) Teacher's guide	SEA-PLM: TC09	3	Substantial refinement: Reduction and modification of the item

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		16.a)Wall chart of any kind 16.b)Working power outlets 16.c)Dictionary 16.d)Class set of textbooks 16.e)Classroom library, book corner or book box 16.f)Bookshelves 16.g)Lockable cabinet 16.h)Teacher desk 16.i)Enough desks for all students 16.j)Working television / monitor 16.k)Working overhead / LCD projector 16.l)Working computer 16.m)Interactive whiteboard Yes. No.	b) Electricity/power c) Working computer available for teaching d) Interactive whiteboard/ smartboard e) Internet available for teaching f) Wall chart of any kind g) Class set of student basic supplies (pencil/pen and exercise book) h) Sufficient chairs or benches i) Sufficient desks or tables Yes No			set (13 items to 9) to focus on infrastructure.
Instructional Resources and Classroom Infrastructure Inventory	B	Do students in your school have textbooks for instruction in <test language>? Yes, every student has at least one. Yes, but not enough. Yes, but so few. No, there are no textbooks.	How does your school make textbooks available to students in this class? a) Mathematics b) <language of test> There are no textbooks for students One textbook is shared by more than two students One textbook is shared by two students Most students have their own textbook, but some have to share Each student has their own textbook	PISA_D:TQ08	3	Structural redesign: Textbook availability measure redesigned for better granularity in sharing ratios.
School- or Classroom library and use	A	18.a) Do you have a library or reading corner in your classroom? 18.b)If Yes, about how many books with different	How many books with different titles are in your classroom library or reading corner? My classroom does not have a library or reading corner	PIRLS:TQR13A,B	3	Significant synthesis: Combined library presence and book quantity into a single

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		titles are in your classroom library? a: Yes. No. b: 0-25 26-50 51-100 More than 100	Fewer than 25 26 to 50 51 to 100 More than 100			complex categorical item.
School- or Classroom library and use	B	Can the students borrow books from the classroom library or reading corner to take home? Yes. No.	Can the students borrow books from the classroom library or reading corner to take home? Yes No	PIRLS:TQR13E	1	Identical measure: Maintained standard PIRLS accessibility (borrowing) measure.
EAS_Instructional and Curricular Leadership (incl. Collaboration & relational school climate)	A	How would you characterize each of the following within your school? a) Teachers' understanding of the school's curricular goals b) Teachers' degree of success in implementing the school's curriculum c) Teachers' expectations for student achievement d) Teachers' ability to inspire students Very high High Medium Low Very Low	How would you characterize each of the following within your school? a) Teachers' understanding of the school's curricular goals b) Teachers' degree of success in implementing the school's curriculum c) Teachers' expectations for student achievement d) Teachers' ability to inspire students Very high High Medium Low Very Low	TIMSS:TQG06 a,b,c,d	1	Direct adoption of standard TIMSS instructional leadership and curricular goal items; properties remain stable.
EAS_Parental Support	A	How would you characterize each of the following within your school? e) Parental involvement in school activities f) Parental commitment to	How would you characterize each of the following within your school? a) Parental involvement in school activities b) Parental commitment to	TIMSS:TQG06 e,f,g,h	1	Direct adoption of standard TIMSS parental support items; parameters remain stable.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		ensure that students are ready to learn g) Parental expectations for student achievement h) Parental support for student achievement Very high High Medium Low Very Low	ensure that students are ready to learn c) Parental expectations for student achievement d) Parental support for student achievement Very high High Medium Low Very Low			
EAS_Parental Support	B	To what extent do you agree or disagree with the following statements 20.a) Parents have a positive attitude towards the school 20.b) Parents care about their child's academic achievement 20.c) Parents meet with teachers to discuss their child's academic achievement 20.d) Parents feel part of the school community Strongly agree Agree Disagree Strongly disagree		SEA-PLM:TC08k-n		Secondary items removed for brevity.
SOS_Personal Safety (including bullying)	A	During the last month, how often did these situations happen at this school? - 1. A student threatened another student - 2. A student hit or hurt another student - 3. A student was discriminated against/excluded by other	How often have these things happened at your school in the past year? a) A student threatened another student b) A student hit or hurt another student c) A student was discriminated against/excluded by other students because of his/her	ERCE:PQ6:Q32.1-4	3	Significant shift: Moved from a general safety opinion to reporting frequencies of specific hostile incidents/bullying.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		students because of his/her appearance, ethnicity, or social class. - 32.4. A student insulted a teacher Never. Sometimes. Many times. Every day or almost every day.	appearance, ethnicity, or social class d) A student insulted a teacher At least once a week Once or twice a month A few times Never			
SOS_Personal Safety (including bullying)	B	Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements. (I feel safe at this school.) Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree or disagree with the following statements about your current school? a) I feel safe at this school Agree a lot Agree a little Disagree a little Disagree a lot	TIMSS:TQG07 a (= PIRLS:TQG11 b)	2	Category standardization: Standardized hostile incident frequency categories to standard TIPI scales.
SOS_Student Behavior & Respect	A	How much do you agree or disagree with the following statements about your current school? c) The students behave in an orderly manner. d) The students are respectful of the teachers e) The students respect school property Agree a lot Agree a little Disagree a little Disagree a lot	How much do you agree or disagree with the following statements about your current school? a) This school's security policies and practices are sufficient b) The students behave in an orderly manner c) The students are respectful of the teachers d) The students respect school property e) This school has clear rules about student conduct f) This school's rules are enforced in a fair and consistent manner Agree a lot Agree a little	TIMSS:TQG07 c-e	3	Significant expansion: Expanded from 3 to 6 items to include security policies and fair enforcement perceptions.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			Disagree a little Disagree a lot			
SOS_Student Behavior & Respect	B	How often do the following situations occur in the <classroom assessed>? 22.a) When the class starts, I have to wait a long time for the students to be silent 22.b) I lose a lot of time because students disrupt the class. 22.c) Students make a lot of noise in the classroom. Never. In some classes. In about half of the classes. In more than half or all classes.	How often do the following situations occur in the <classroom assessed>? a) When the class starts, I have to wait a long time for the students to be silent b) I lose a lot of time because students disrupt the class c) Students make a lot of noise in the classroom Very often Often Sometimes Never	ERCE:PQ6:Q25.1-3	2	Standardized categories: Response categories adjusted to match standard TIPI frequency scales.
External Supervision & Inspection	A	Since the beginning of the year, have you seen a school inspector (inspector or educational adviser) in your class? Yes. No.	Since the beginning of this school year, how often have the following people visited your classroom or discussed your teaching with you (e.g., through observation, feedback, or pedagogical meetings)? a) External supervisors (e.g., school inspector, educational adviser, pedagogical advisor) b) School management (e.g., principal, deputy principal, pedagogical coordinator) c) Internal pedagogical staff (e.g., mentor, senior teacher, subject coordinator) At least once a week Once or twice a month A few times a year Never	PASEC:QM_43	3	Significant shift: Moved from a binary "Yes/No" check to a frequency scale covering multiple personnel roles.
External Supervision & Inspection	B	How often does the management team of this school or any of its		ERCE:PQ6:Q33.1-7		No secondary question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		members perform the following activities? 23.a) Visit our classrooms and see how we teach. 23.b) Gives us feedback on the way we teach. 23.c) Gives us feedback on how we assess students. 23.d) It provides us with feedback on our planning. 23.e) Gives us feedback on our handling of the course group. 23.f) Organizes or promotes instances of coordination and collaboration among teachers. 23.g) Organizes meetings with all teachers to provide pedagogical feedback. Never Once per year Two or three times a year One or more times a month A few times per week or every day				
Class size/ number of students on day of test	A	A. How many students are in this class? B. How many of the students in #xA are in the <target grade>? _____ students (Write in the number.)	How many students are in this class? a) How many students are in this class in total? b) How many of these students are in the <target grade>? Male students (Write in the number.) Female students (Write in the number.)	TIMSS:TQG10A,B	3	Significant shift: Transformed a single numeric student total into gender-segregated student counts.
Class size/ number of students on day of test	B	Do you teach <multi-grade classrooms> in this school? Yes. No.		PISA_D:TQ05		No secondary question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Teaching conditions/ Teaching Constraints	A	In your view, to what extent do the following limit how you teach this class? a) Students lacking prerequisite knowledge or skills b) Students suffering from lack of basic nutrition c) Students suffering from not enough sleep d) Students absent from class e) Disruptive students f) Uninterested students g) Distracted students h) Students with mental, emotional, or psychological impairment i) Students with difficulties understanding the language of instruction Not at all Some A lot	In your view, to what extent do the following limit how you teach this class? a) Students lacking prerequisite knowledge or skills b) Students suffering from lack of basic nutrition c) Students suffering from not enough sleep d) Students absent from class e) Disruptive students f) Uninterested students g) Distracted students h) Students with mental, emotional, or psychological impairment i) Students with difficulties understanding the language of instruction Not at all Some A lot	TIMSS:TQG13 a-i	1	Direct adoption of the 9-item TIMSS teaching constraints scale; parameters remain stable.
Teaching conditions/ Teaching Constraints	B	The demands of parents and guardians hinder my pedagogical work. Strongly disagree. Disagree. Agreed. Very much in agreement.	To what extent do the following limit how you teach this class? a) Lack of basic resources or equipment (e.g., textbooks, furniture, electricity, or computers and internet access) b) Poor school infrastructure c) Safety or security concerns in or around the school d) Teaching several grade levels in the same class Not at all Some A lot	ERCE:PQ6:Q28.5	3	Significant construct shift: Shifted from parental pressure to infrastructural and safety constraints.
Language of instruction	A	What language do you use when teaching your	What languages do you use when teaching your students in	PISA_D:TQ23	1	Identical PISA-D scale; minor pluralization of

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		students in this class? Only <language of instruction>. Mostly <language of instruction> but sometimes their home language. Sometimes <language of instruction> and sometimes their home language. Mostly their home language. Always their home language.	this class? Only <language of instruction> Mostly <language of instruction> but sometimes students' home languages Sometimes <language of instruction> and sometimes students' home languages Mostly students' home languages Always students' home languages			"languages" to better fit multilingual classroom contexts.
Instructional time (Language/ Reading or mathematics)	A	In a typical week, how much time do you spend teaching mathematics to the students in this class? _____ minutes per week (Please convert the number of hours into minutes.)	In a typical week, how much time do you spend teaching mathematics to the students in this class? _____ minutes per week (Please convert the number of hours into minutes.)	TIMSS:TQM01	1	Identical numeric recall: Maintained minute-based recall for Mathematics instruction time.
Instructional time (Language/ Reading or mathematics)	B	In a typical week, how much time do you spend on <language of test> language instruction and/or activities with the students? _____ minutes per week	In a typical week, how much time do you spend on <language of test> language instruction and/or activities with the students? _____ minutes per week	PIRLS:TQR04	1	Identical numeric recall: Maintained minute-based recall for Language instruction time.
Instructional activities (general, Language/ Reading, Math)	A	How often do you do the following in teaching this class? a. Relate the lesson to students' daily lives b. Ask students to explain their answers c. Communicate goals or objectives for the lesson to the students d. Ask students to complete challenging exercises that require them to go	How often do you do the following in teaching this class? a) Relate the lesson to students' daily lives b) Ask students to explain their answers c) Communicate goals or objectives for the lesson to the students d) Ask students to complete challenging exercises that require them to go beyond the instruction	TIMSS:TQG12 a-g	2	Item set expansion: Expanded from 7 to 11 items to include tests and differentiation while stem/scale remain stable.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		beyond the instruction e. Encourage classroom discussions among students f. Link new content to students' prior knowledge g. Ask students to decide their own problem solving procedures Every or almost every lesson About half the lessons Some lessons Never	e) Encourage classroom discussions among students f) Link new content to students' prior knowledge g) Ask students to decide their own problem solving procedures h) Check or review students' homework or classwork i) Take a written quiz or test j) Perform differentiated activities according to abilities k) Have students memorize or repeat important facts, rules, or examples Every or almost every lesson About half the lessons Some lessons Never			

Table 8-7: School Questionnaire Selected Items and Modifications

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Gender	A	Are you female or male? Female Male	Which of these describes you? Female Male <Other>	SEA-PLM, SC01	2	Stem refined for inclusivity and addition of "<Other>" category shifts the binary response scheme; thresholds may shift.
Age	A	What is your age? 24 years or younger 25-34 years 35-44 years 45-54 years 55-64 years 65 years or older	How old are you? Under 25 25-29 30-39 40-49 50-59 60 or more	SEA-PLM, SC02	1	Minor wording update and slight refinement of age bands; core construct and parameters remain stable.
Years of service, experience in education	A	By the end of this school year, how many years will you have	By the end of this school year, how many years will you have	TIMSS: SQ-17/PIRLS: SQ-16	1	Negligible change; maintained original construct of school-specific stability.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		been a principal at this school? ____ years	been a principal altogether? ____ years			
Years of service, experience in education	B	By the end of this school year, how many years will you have been a principal at this school? ____ years	By the end of this school year, how many years will you have been a principal at this school? ____ years	TIMSS SQ-17/PIRLS SQ-16	1	Negligible change; maintained the secondary measure of school-specific tenure.
Highest level of education	A	What is the highest level of formal education you have completed? did not complete ISCED 6 ISCED 6 ISCED 7 ISCED 8	What is the highest level of formal education you have completed? Did not complete <Upper secondary education—ISCED Level 3> <Upper secondary education—ISCED Level 3> <Post-secondary, non-tertiary education—ISCED Level 4> <Short-cycle tertiary education—ISCED Level 5> <Bachelor's or equivalent level—ISCED Level 6> <Master's or equivalent level—ISCED Level 7> <Doctoral or equivalent level—ISCED Level 8>	TIMSS: SQ-19/ PIRLS: SQ-17	2	Expansion of ISCED levels to lower categories (3–5) requires threshold recalibration for LMIC contexts.
Professional development	A	Have you received additional training in your capacity as head teacher over the past two years (seminar, retraining...)? IF yes, please specify the total length Yes No	Have you completed any formal professional training in education administration or management (e.g., leadership or inspection)? Yes No	PASEC: SQ-5,6	3	Significant shift: Synthesis of ERCE/PASEC and removal of the “number of days” duration requirement; full validation required.
Professional development	B	In the last 2 years, have you completed refresher, specialization or postgraduate studies in education administration or management (e.g., leadership, inspection, management)?		ERCE: SQ-15		No second question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		Yes No				
School ownership (private/public/religious)	A	Which of the following best describes your school? Public (<managed directly or indirectly by a public education authority...>) Private (<managed directly or indirectly by a non-government organization...>)	Which of these options best describes your school? Public: managed directly or indirectly by a public education authority, government agency, or governing board appointed by government or elected by public franchise. Private/ Independent: managed directly or indirectly by a non-government organization, e.g., a church, trade union, business, or other private institution.	SEA-PLM, SC-05	3	Significant shift: Merged SEA-PLM and PISA-D to create multi-layered public/private management descriptions.
School ownership (private/public/religious)	B	What kind of organization runs your school? A church or other religious organization Another not-for-profit organization A for-profit organization The government	What kind of organization manages your school? A church or other religious organization A not-for-profit organization A for-profit organization The government Other	PISA-D, Q07	3	NEW dimension: Addition of "Other" and specific categories to capture full religious/NGO scope.
School type/structure (primary/secondary/language)	A	Which of the following levels of education does your school provide? ISCED0 ISCED1 ISCED2 ISCED3	Which of the following levels of education does your school provide? <Early childhood education—ISCED Level 0> <Primary education—ISCED Level 1> <Lower secondary education—ISCED Level 2> <Upper secondary education—ISCED Level 3>	PISA-D, Q02	1	Negligible change; standardized ISCED levels (0–3) maintained from PISA-D.
School location (urban/rural)	A	How many people live in the city, town, or area where your school is located? More than 500,000 people 100,001–500,000 people	How many people live in the city, town, or area where your school is located? More than 500,000 people 100,001–500,000 people	TIMSS/PIRLS, 5A	1	Negligible change; direct adoption of standardized TIMSS/PIRLS population size ranges.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		50,001–100,000 people 30,001–50,000 people 15,001–30,000 people 3,001–15,000 people 3,000 people or fewer	50,001–100,000 people 30,001–50,000 people 15,001–30,000 people 3,001–15,000 people 3,000 people or fewer			
School location (urban/rural)	B	Which best describes the immediate area in which your school is located? Urban—Densely populated Suburban—On fringe or outskirts of urban area Medium size city or large town Small town or village Remote rural	Which best describes the immediate area in which your school is located? Urban—Densely populated Suburban—On fringe or outskirts of urban area Medium size city or large town Small town or village Remote rural	LaNA, SCQ-3	1	Negligible change; maintained established LaNA/TIMSS area categories.
SES Backgrounds of students	A	Approximately what percentage of students in your school have the following backgrounds? a) Come from economically disadvantaged homes/ b) Come from economically affluent homes 0 to 10% 11 to 25% 26 to 50% More than 50%	Approximately what percentage of students in your school have the following backgrounds? a) Come from economically disadvantaged homes b) Come from economically affluent homes 0–10% 11–25% 26–50% 51–75% More than 75%	PIRLS/TIMSS, 3a,b	2	Category expansion: Added higher-end percentage categories (75%+) to capture granular differences in LMIC contexts.
School size (total enrollment of boys and girls and at target grade)	A	What is the total enrollment of students in your school as of <first day of month TIMSS testing begins, 2023>? _____ students	What is the total enrollment of students in your school as of <first day of month testing begins, 202x>? (write “0” if there are none) _____ boys _____ girls	TIMSS/PIRLS, 1	3	Significant shift: Moved from a single numeric total to gender-segregated student counts; recalibration required.
School size (total enrollment of boys and girls and at target grade)	B	As of <August 1, 2017>, what was the total school enrolment (number of students)? Number of boys: _____ Number of girls: _____		PISA-D, Q03		No second question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Enrollment & number of classes in target grade	A	What is the total enrollment of <fourth grade> students in your school as of <first day of month TIMSS testing begins, 2023>? ____ students	What is the total enrollment of <target grade> students in your school as of <first day of month testing begins, 202x>? (write "0" if there are none) ____ boys ____ girls	TIMSS/PIRLS, 2	3	Significant shift: Moved to require gender-segregated data specifically for the target grade.
Language of test/ language of instruction	A	Approximately what percentage of students in your school have <language of test> as their native language? More than 90% 76 to 90% 51 to 75% 26 to 50% 25% or less	Approximately what percentage of students in your school have <language of test> as their first language at home? More than 90% 76–90% 51–75% 26–50% 25% or less	TIMSS: SQ-4	1	Minor wording update for precision ("first language") while maintaining identical categories.
Number of teachers/ Staff	A	What is the total current number of teachers at your school? ____ full time teachers ____ part time teachers	How many teachers are at your school? (A full-time teacher is employed at least 90% of the time as a teacher for the full school year.) ____ full time teachers (> 90% of ull-time hours) ____ part time teachers (71-90% of full-time hours) ____ part time teachers (50-70% of full-time hours) ____ part time teachers (less than 50% of full-time hours)	LaNA, Q2	3	Significant shift: Expanded categorical time-use brackets for part-time staff based on TALIS/STEM definitions.
School Educational Resources except books (and affection by shortage)	A	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? b) Supplies (e.g., papers, pencils, materials) Not at all A little Some A lot	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? a) Educational resource materials (e.g., papers, pencils, materials) b) Resources for students with learning disabilities Not at all	LaNA, Q10a	2	New dimension added: Included an explicit check for resources for students with learning disabilities.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			A little Some A lot			
School human resources (and affected by shortage)	A	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? c) Qualified teachers To a large extent To a moderate extent To a little extent Not at all	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? a) Qualified teachers Not at all A little Some A lot	LaNA, Q10c	2	Scale standardization: Standardized the response scale (4-point) to match other resource shortage items.
School Library (equipment and use)	A	Does your school have a school library? Yes No	Approximately, how many books does your school library have? My school does not have a school library 50 books or fewer 51 to 100 books 101 to 200 books 201 to 500 books 501 to 1000 books 1001 books or more	PIRLS, SCQ-7A	3	Significant synthesis: Combined existence and quantity into one complex categorical item with new ranges.
School Library (equipment and use)	B	How many books does your library have? Less than 100 books, Between 100 and 200 books, Between 200 and 500 books, Between 500 and 1000 books, 1000 books or more		SEA-PLM, SC19		No second question provided.
Internet	A	Do students in your school have access to these resources? B) High speed internet Yes No	Does your school have the following resources? a) Reliable internet connection available for teaching and learning Yes No	TIMSS, SCQ-10	2	Refined focus: Shifted from general high-speed access to reliable access for teaching and learning.
School facilities (incl. Toilets SDG)	A	Does your school have the following resources? a) Running water b) Electricity	Does your school have the following resources? a) Electricity b) Safe drinking water	Lana, SQ-6	3	Significant expansion: Massive addition of items (handwashing, disability)

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		c) Suitable toilets or latrines d) Playground e) Computers or tablets for students Yes No	available to students c) Handwashing facilities with soap and water d) Facilities accessible to students with disabilities (e.g., ramps, adapted toilets) e) Availability of toilets or latrines that are separate for girls and boys f) Sufficient digital devices (computers, laptops, etc.) for school administrative purposes g) Sufficient digital devices (computers, laptops, etc.) for teaching Yes No			access, separate toilets) to meet SDG standards.
School facilities (incl. Toilets SDG)	B	Which of these services does the school have? 1.) Electric light 2.) Drinking water 3.) Drainage or sewerage 4.) Telephone 5.) Mobile computer laboratory 6.) Restrooms in good condition 7.) Internet connection 8.) Garbage collection 9.) Student transportation (free of charge) Yes No		ERCE-33.1-9		No second question provided.
Classroom facilities	A	Does your school have the following classroom resources a) Chairs or benches b) desks or tables c) textbooks for each students Yes No	Does your school have the following classroom resources? a) Sufficient chairs or benches b) Sufficient desks or tables c) Cupboards d) Bookshelves e) Chalkboard or a white board f) Smart board	LaNA, SQ-7	3	Significant shift: Added multiple teaching aid categories (Smart boards, projectors) and a "Sufficient" qualifier.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
			g) Digital projector Yes No			
Availability of textbooks	A	How many textbooks for each student does your school have for <Grade 5> mathematics...? No text books are available for students One text book for each student Two students share one textbook More than two students share a textbook	How does your school make textbooks available to students in <target grade>? a) Mathematics b) <language of test> There are no textbooks for students One textbook is shared by more than two students One textbook is shared by two students Most students have their own textbook, but some have to share Each student has their own textbook	SEA-PLM, SC16	2	Refined response scheme to capture nuanced sharing scenarios (1:2 vs. shared by more than two).
Availability of textbooks	B	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? a) Instructional materials (e.g., textbooks) To a large extent To a moderate extent To a little extent Not at all	How much is your school's capacity to provide instruction affected by a shortage or inadequacy of the following? a) Instructional materials (e.g., textbooks) Not at all A little Some A lot	LaNA, 10a	2	Standardized scale: Shortage response options moved to the uniform 4-point scale.
Provision of additional support in reading or math	A	Additional support? Yes, for free Yes, partially or fully paid by parents No My school does not teach this [grade]	Does your school provide extra classes or tutoring for students who need additional support in learning? a) In reading b) In mathematics Yes, at no additional cost to families Yes, for a fee No	AMPL: SC13A	2	Clarified wording and split tutoring into subject-specific tracks (Reading/Math).

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
Digital learning resources	A	Does your school provide students with access to digital learning resources (e.g., books, videos)? Yes No	How many digital devices does your school have for use by <target grade> students? ____ devices	PIRLS, SQ-8	3	Significant format change: Changed from broad binary access to a specific numeric count of devices.
Digital learning resources	B	How many computers (including tables) does your school have for use by <target grade> students? ____ computers		PIRLS, SQ-09		No second question provided.
School Year Structure (beginning/end of school year, operational days)	A	For the <fourth grade> students in your school: How many days per year is your school scheduled to be open for instruction? Write in number	For the <target grade> students in your school: How many days per year is your school scheduled to be open for instruction? ____ days	LaNA, SCQ-5a/PIRLS/TIMSS, SCQ-6A	1	Minor stylistic adjustment; core numeric recall for days open remains the same.
School Year Structure (beginning/end of school year, operational days)	B	In one calendar week, how many days is the school open for instruction? 6 days 5 1/2 days 5 days 4 1/2 days 4 days Other	In one calendar week, how many days is the school open for instruction? 6 days 5 1/2 days 5 days 4 1/2 days 4 days Other	LaNA, SCQ-5d	1	Identical measure: Secondary weekly instruction source remains identical.
School Year Structure (instructional time per day)	A	What is the total instructional time, excluding breaks, in a typical day? ____ minutes	What is the total instructional time, excluding breaks, on a typical day? ____ minutes	LaNA, SCQ-5B, PIRLS, SCQ-6B	1	Negligible wording adjustment; numeric minute input maintained.
Shifts	A	Does your school have double shift or double session? (...) Yes No	Does your school have double shifts or double sessions? (one group of students attends in the morning and another in the afternoon) Yes No	PASEC: QD_23	1	Minor wording expansion for clarity; construct of double shifts is identical.
Shifts	B	Regarding the operation of this school building, what type of school day does the school have? Morning shift only It has a double shift (one group of		ERCE: SQ-28		No second question provided.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		students attends in the morning and another in the afternoon) It has a full day (morning and afternoon) It has a different type of day				
Days school closed (Covid, strikes, etc.)	A	Within the last 12 months, about how many days was your school closed due to unforeseen circumstances (e.g., strikes, bad weather, etc.)? _____ days	Within the last 12 month, about how many days was your school closed due to unforeseen circumstances (e.g., strikes, bad weather, etc.)? _____ days	LaNA, SCQ-5b	1	Identical measure: Maintained the original source format regarding unforeseen closures.
Teacher job satisfaction	A	How would you characterize each of the following within your school? a) Teachers' job satisfaction Very high High Medium Low Very low	How would you characterize each of the following within your school? a) Teachers' job satisfaction Very high High Medium Low Very low	LaNA, SCQ-11a	1	Standardization: Characterization scale maintained from the original LaNA source.
EAS_Instructional and Curricular Leadership and Staff Cohesion	A	How would you characterize each of the following within your school? b) Teacher's understanding of the school's curricular goals c) Teacher's degree of success in implementing the school's curriculum d) Teacher's expectations for school achievement Very high High Medium Low Very low	How would you characterize each of the following within your school? a) Teacher's understanding of the school's curricular goals b) Teacher's degree of success in implementing the school's curriculum c) Teacher's expectations for school achievement Very high High Medium Low Very low	LaNA, SCQ-11b-d	1	Direct adoption of standard characterization scale for curricular goals and expectations.
EAS_Parental Support	A	How would you characterize each of the following within your school? e) Parental support for student	How would you characterize each of the following within your school? a) Parental support for student	LaNA, SCQ-11e-f	1	Direct adoption of the perceived parental support and involvement measure.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		achievement f) Parental Involvement in school activities Very high High Medium Low Very low	achievement b) Parental involvement in school activities Very high High Medium Low Very low			
EAS_Student Engagement and Peer Culture	A	How would you characterize each of the following within your school? g) Student's regard for school property h) Student's desire to do well in school Very high High Medium Low Very low	How would you characterize each of the following within your school? a) Students' regard for school property b) Students' desire to do well in school Very high High Medium Low Very low	LaNA, SCQ-11g-h	1	Standard scale maintained for student regard for property and desire to succeed.
SOS_Personal Safety / BULLYING (incl. Teachers)	A	To what degree is each of the following a problem among <fourth grade> students in your school? h) Intimidation or verbal abuse among students (including texting, emailing, etc.) i) Physical fights among students j) Intimidation or verbal abuse of teachers or staff (including texting, emailing, etc.) Not a problem Minor problem Moderate problem Serious problem	To what degree is each of the following a problem among <target grade> students in your school? a) Intimidation or verbal abuse among students (including texting, emailing, etc.) b) Physical fights among students c) Intimidation or verbal abuse of teachers or staff, including online harassment (e.g., messaging, social media, or sharing images or videos online) Not a problem Minor problem Moderate problem Serious problem	LaNA, SCQ-12h-j	2	Refinement: Expanded item (c) to explicitly include online harassment and social media examples.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
SOS_School rules and enforcement	A	To what degree is each of the following a problem among <fourth grade> students in your school? a) Arriving late at school b) Absenteeism (i.e., unjustified absences) Not a problem Minor problem Moderate problem Serious problem	To what degree is each of the following a problem among <target grade> students in your school? a) Arriving late at school b) Absenteeism (i.e., unjustified absences) Not a problem Minor problem Moderate problem Serious problem	LaNA, SCQ-12a-b	1	Standard scale maintained for the severity of lateness and absenteeism problems.
SOS_Student Behavior & Respect	A	To what degree is each of the following a problem among <fourth grade> students in your school? c) Classroom disturbance d) Cheating e) Profanity f) Vandalism g) Theft Not a problem Minor problem Moderate problem Serious problem	To what degree is each of the following a problem among <target grade> students in your school? a) Classroom disturbance b) Cheating c) Profanity d) Vandalism e) Theft Not a problem Minor problem Moderate problem Serious problem	LaNA, SCQ-12c-g	1	Standard scale maintained for cheating, profanity, and vandalism.
Policies on grade retention/ academic support & digital devices	A	Which policies regarding <grade repetition> are implemented in your school? a) <Grade repetition> is possible on a voluntary basis, i.e., with request or permission from the parents b) If a student fails the minimum achievement standards at the end of the school year, he or she has to repeat the grade. c) Individual courses can be repeated without repeating a whole grade. d) After repeating grades a certain number of times, students are	What are your school's policies regarding <grade repetition>? a) Students may <repeat a grade> voluntarily (e.g., at the request of or with permission from parents) b) Students must <repeat a grade> if they do not meet minimum achievement standards at the end of the school year c) Students may <repeat individual courses> without repeating the entire <grade> d) Students are expected to leave the school after	PISA: Q23	1	Negligible wording adjustment; original PISA policy list and Yes/No options maintained.

Indicator	Part	Original Item & Options	Suggested Item & Options	Source(s)	Mod. Level	Justification
		expected to leave the school. e) <Grade repetition> is prohibited by <state or district> regulations. f) <Grade repetition> is prohibited by school policies. Yes No	<repeating grades> a certain number of times e) <State or district> regulations prohibit <grade repetition> f) School policies prohibit <grade repetition> Yes No			
Teacher Absenteeism, Teacher availability, Issues at school	A	To what degree is each of the following a problem among teachers in your school? a) Arriving late or leaving early b) Absenteeism Not a problem Minor problem Moderate problem Serious problem	To what degree is each of the following a problem among teachers in your school? a) Arriving late or leaving early b) Absenteeism c) Serious behavioral misconduct by teachers or staff (e.g., substance use, inappropriate behavior towards students or colleagues) Not a problem Minor problem Moderate problem Serious problem	LaNA, SCQ-13a,b	2	New dimension added for behavioral misconduct (e.g., substance use or inappropriate behavior).
Nutrition, meals provided by school	A	Does your school provide free meals for students? a) breakfast b) Lunch Yes, for all students Yes, for some students No	Does your school provide free meals for students? a) Breakfast b) Lunch Yes, for all students Yes, for some students No	LaNA, SCQ-4	1	Standard check maintained for school-provided meals (Breakfast/Lunch).

8.7 Appendix A7: AI Diligence Statement

In developing the R&D project on improving IEA questionnaires for use in low- and middle-income countries, the project team used AI tools, including Consensus, ChatGPT, and Claude, to support selected stages of the work. The project aimed to enhance the contextual relevance and validity of background questionnaire information for LMIC contexts through the review of existing literature, conceptual frameworks, regional and international assessment instruments, and psychometric evidence from available assessment data.

8.7.1 Creation Diligence

AI tools were used as support instruments throughout the creation process, but not as autonomous decision-makers. Their contributions included support for literature review activities, identification and summarization of relevant sources, extraction of key points from literature and assessment documentation, generation of ideas for developing and refining item-suitability and adaptation-classification schemes and drafting or revising selected text elements.

AI assistance was used across three main types of activity: exploratory support, including idea generation and literature exploration; technical assistance, including the development and review of scripts for statistical analyses, combination of information from multiple data sources, and generation and refinement of graphics; and drafting support, including the preparation and revision of selected text elements and sections related to data interpretation.

AI tools were also used to review and cross-check selected elements of manual work related to framework development, indicator classification, and the matching of questionnaire items to indicators.

Materials processed with AI tools included publicly available questionnaires and codebooks, assessment data available for research use from completed regional and international assessment cycles, and project-internal indicator lists, classification schemes, draft text, and analysis specifications. No directly identifiable respondent information was used in AI-supported work.

8.7.2 Transparency Diligence

The intended audiences for the project include researchers, questionnaire and test developers working in LMIC contexts, IEA management, and other stakeholders interested in improving contextual data collection in international and regional learning assessments.

Given these audiences, transparency about AI use is important. AI contributed to several stages of the work, including idea generation, structuring of classification approaches, literature support, summarization, drafting, programming assistance, graphics preparation, and verification of selected manual and analytical steps. These contributions were treated as support to expert work rather than as a substitute for professional judgment.

All AI-generated or AI-supported outputs were reviewed, revised, and integrated by the project team. Where AI outputs informed decisions, they were checked against source materials, statistical results, project objectives, and expert knowledge of international large-scale assessments and LMIC questionnaire development., and expert understanding of international large-scale assessments and LMIC questionnaire development.

8.7.3 Deployment Diligence

Several safeguards were applied to verify the accuracy, appropriateness, and usefulness of AI-supported outputs. Statistical results were checked through recalculation of selected examples and plausibility checks against the underlying data and related graphical outputs. AI-generated graphics and summaries were reviewed in relation to the broader analytical findings. Draft text produced or revised with AI support was reworked, reviewed, and checked by project team members.

For classification tasks, AI-supported suggestions were not accepted automatically. They were reviewed against the agreed classification scheme, compared with the original questionnaire materials, and, where necessary, subjected to human arbitration. In selected cases, separate AI tools were used to challenge or verify outputs generated by another tool.

All substantive decisions, including methodological choices, classification rules, and interpretation of results, were made by the project team. Final outputs were reviewed by multiple team members, with attention to consistency across report sections, questionnaire materials, indicators, statistical evidence, and methodological explanations. The project team retains full responsibility for the final product, including its accuracy, interpretation, presentation, and suitability for the intended audience.

References

- Abadzi, H. (2007). *Absenteeism and Beyond: Instructional Time Loss and Consequences*. Retrieved from <https://hdl.handle.net/10986/7569>
- Australian Council for Educational Research (2022). *Minimum proficiency levels unpacked* (WG/GAML/4). Retrieved from UNESCO Institute for Statistics website: https://www.uis.unesco.org/sites/default/files/medias/fichiers/2025/08/WG_GAML_4_MPLs-Unpacked_ACER.pdf
- Baker, David, P., Goesling, B., & Letendre, Gerald, K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the “Heyneman-Loxley effect” on mathematics and science achievement. *Education Policy Studies*, 46(3), 291–312. Retrieved from <https://doi.org/10.1086/341159>
- Bendini, M., & Devercelli, A. E. (2022). *Quality Early Learning: Nurturing Children’s Potential*.
- Blömeke, S., Nilsen, T., Olsen, R.V., Gustafsson, JE. (2022). *Conceptual and Methodological Accomplishments of ILSAs, Remaining Criticism and Limitations*. In: Nilsen, T., Stancel-Piątak, A., Gustafsson, JE. (eds) *International Handbook of Comparative Large-Scale Studies in Education*.: Springer International Handbooks of Education.
- Bold, T., Filmer, D., Martin, G., Molina, E., Rockmore, C., Stacy, B., . . . Wane, W. (2017). *What Do Teachers Know and Do? Does It Matter? Evidence from Primary Schools in Africa*. Policy Research Working Paper.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371–399. <https://doi.org/10.1146/annurev.psych.53.100901.135233>
- Brese, F., & Mirazchiyski, P. (2013). *Issues and methodologies in large-scale assessments: Special issue 2: Measuring students family background in large scale international education studies*. IERI monograph series.
- Carroll, J. B. (1963). A Model of School Learning. *Teachers College Record: The Voice of Scholarship in Education*, 64(8), 1–9. <https://doi.org/10.1177/016146816306400801>
- Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., & Rogers, F. H. (2006). Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1), 91–116. Retrieved from <https://pubs.aeaweb.org/doi/pdfplus/10.1257/089533006776526058>
- Chudgar, A., & Luschei, T. (2009). National Income, Income Inequality, and the Importance of Schools: A Hierarchical Cross-National Comparison. *American Educational Research Journal*, 46(3), 626–658. Retrieved from <https://doi.org/10.3102/0002831209340043>
- Confraria, H., Ciarli, T., & Noyons, E. (2024). Countries’ research priorities in relation to the Sustainable Development Goals. *Research Policy*, 53(3), 104950. <https://doi.org/10.1016/j.respol.2023.104950>
- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, 76(1), 1–62. Retrieved from <https://doi.org/10.3102/00346543076001001>
- Creemers, B., & Kyriakides, L. (2008). The Dynamics of Educational Effectiveness. Advance online publication. <https://doi.org/10.4324/9780203939185>

- D'Adamo, I., Della Sciucca, M., Gastaldi, M., & Lupi, B. (2025). Indicator Assessment of Sustainable Development Goals: A Global Perspective. *Sustainability*, 17(18), 8259. <https://doi.org/10.3390/su17188259>
- David, N., Temba, J., & Matondo, S. S. (2025). Reasons behind Long Commute Distances on Students' Attendance and Academic Performance in Mwanza City, Tanzania. *International Journal of Teaching, Learning and Education*, 4(3), 97–102. <https://doi.org/10.22161/ijtle.4.3.11>
- Davies, M. von, & Kennedy, A. (Eds.) (2025). *TIMSS 2027 Assessment Frameworks: TIMSS & PIRLS International Study Center*, Boston College.
- Desforges, C., & Abouchaar, A. (2003). *The Impact of Parental Involvement, Parental Support and Family Education on Pupil The Impact of Parental Involvement, Parental Support and Family Education on Pupil Achievements and Adjustment: A Literature Review*. Retrieved from https://www.nationalnumeracy.org.uk/sites/default/files/documents/impact_of_parental_involvement/the_impact_of_parental_involvement.pdf
- Dolata, S. (2008). Indice du statut socioéconomique du milieu familial des élèves du SACMEQ. *Mesure et évaluation en éducation*, 31(1), 121–149. <https://doi.org/10.7202/1025015ar>
- Evans, D., & Yuan, F. (2018). *The Working Conditions of Teachers in Low- and Middle-Income Countries*. Retrieved from <https://riseprogramme.org/sites/default/files/inline-files/Yuan.pdf>
- Flannery, D., Palcic, D., & Clavel, J. (2024). *Does Homework Improve Student Achievement? Evidence from Developed and Developing Countries*. Retrieved from <https://ssrn.com/abstract=4940700>
- Friedman, T., Schwantner, U., Spink, J., Tabata, N., & Waters, C. (2016). *Improving Quality Education and Children's Learning Outcomes and Effective Practices in the Eastern and Southern Africa Region: Main report*.
- Fuller, B. (1987). What School Factors Raise Achievement in the Third World? *Review of Educational Research*, 57(3), 255–292. <https://doi.org/10.3102/00346543057003255>
- Goos, M., Pipa, J., & Peixoto, F. (2021). Effectiveness of grade retention: A systematic review and meta-analysis. *Educational research review*, 34(1), 100401. <https://doi.org/10.1016/j.edurev.2021.100401>
- Grisay, A. (1984). Analyse Des Inégalités De Rendement Liées Au Sexe De l'élève Dans l'Enseignement Primaire Ivoirien. *International Review of Education*, 30(1), 25–39. <https://doi.org/10.1007/BF00598016>
- Grisay, A., & Griffin, P. Measuring educational progress through assessment. In *The International Encyclopedia of Education*.
- Gubbels, J., van der Put, C. E., & Assink, M. (2019). Risk Factors for School Absenteeism and Dropout: A Meta-Analytic Review. *Journal of Youth and Adolescence*, 48(9), 1637–1667. <https://doi.org/10.1007/s10964-019-01072-5>
- Hanushek, E. A. (1995). INTERPRETING RECENT RESEARCH ON SCHOOLING IN DEVELOPING COUNTRIES. *The World Bank Research Observer*, 10(2), 227–246. <https://doi.org/10.1093/wbro/10.2.227>

- Hastedt, D., Sibberns, H. (2022) (2022). *Future Directions, Recommendations, and Potential Developments of ILSA*. In: Nilsen, T., Stancel-Piątak, A., Gustafsson, JE. (eds) *International Handbook of Comparative Large-Scale Studies in Education*: Springer International Handbooks of Education.
- Hattie, J. A. C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*: Routledge.
- Heyneman, S. P., & Loxley, W. A. (1983). The Effect of Primary-School Quality on Academic Achievement Across Twenty-nine High- and Low-Income Countries. *American Journal of Sociology*. Advance online publication. <https://doi.org/10.1086/227799>
- Heyneman, S., & Lee, B. (2013). Impact of International Studies of Academic Achievement on Policy and Research. In L. Rutkowski, M. Davier, & D. Rutkowski (Eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues, and Methods of Data Analysis* (pp. 37–75). Chapman and Hall/CRC.
- Hochweber, J., & Vieluf, S. (2018). Gender differences in reading achievement and enjoyment of reading. The role of perceived teaching quality. *The Journal of Educational Research*, 111, 268–283.
- Holland, P. W., & Wainer, H. (2012). *Differential Item Functioning*: Routledge.
- Howie, S. J. (2022). *Regional Studies in Non-Western Countries, and the Case of SACMEQ, 2022*. in T Nilsen, A. Stancel-Piątak, & J-E, Gustafsson (Eds) (2022). *International Handbook of Comparative Large-Scale Studies in Education: Perspectives, Methods and Findings*: Springer.
- International Labour Office, & United Nations Children’s Fund (2021). *Child Labour: Global estimates 2020: Trends and the road forward*. New York.
- Kaplan, D., Jude, N., Harra, K., & Stampka, J. (2025). *On the Development of Probabilistic Projections of Country-level Progress to the UN SDG Indicator of Minimum Proficiency in Reading and Mathematics*. Preprint (No. arXiv:2511.06107). Retrieved from arXiv website: <https://doi.org/10.48550/arXiv.2511.06107>
- Karamperidou, D., Brossard, M., Peirollo, S., & Richardson, D. (2020). *Time to Teach: Teacher attendance and time on tasks in Eastern and Southern Africa*.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*: Guilford Publications.
- Kuger, S., Klieme, E., Jude, N., & Kaplan, D. (2016). *Assessing Contexts of Learning*. Cham: Springer International Publishing.
- Li, X. (2021). *PISA and PISA for ‘development’: an inquiry into the OECD’s expansion into low- and middle-income countries* (Doctoral dissertation), University College London. UCL Discovery.
- Liu, Z., Zhang, W., & Hu, L. (2024). Can desktop virtual reality effectively enhance academic achievement? —A meta-analysis. *Innovations in Education and Teaching International*. Advance online publication. <https://doi.org/10.1080/14703297.2024.2354753>
- McDonnell, L. M. (1995). Opportunity to Learn as a Research Concept and a Policy Instrument. *Educational Evaluation and Policy Analysis*, 17(3), 305–322. <https://doi.org/10.3102/01623737017003305>

- Mejía-Rodríguez, A. M., & Kyriakides, L. (2022). What matters for student learning outcomes? A systematic review of studies exploring system-level factors of educational effectiveness. *Review of Education*, 10(3), 89. <https://doi.org/10.1002/rev3.3374>
- Mislevy, R. J. (1991). Randomization-Based Inference about Latent Variables from Complex Samples. *Psychometrika*, 56(2), 177–196. <https://doi.org/10.1007/BF02294457>
- Mullis, I. V. S., Martin, M. O., & von Davier, M. (Eds.) (2021). *TIMSS 2023 Assessment Frameworks*: TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College, and International Association for the Evaluation of Educational Achievement (IEA).
- OECD. *PISA 2015 Results (Volume III): Students' Well-Being*.
- OECD (2018). *PISA for Development Assessment and Analytical Framework: Reading, Mathematics and Science*. PISA: OECD Publishing, Paris.
- OECD (2026). *PISA 2029 media and artificial intelligence (MAIL) assessment framework: First draft*. Retrieved from [https://www.oecd.org/content/dam/oecd/en/about/projects/edu/the-pisa-2029-media-and-artificial-intelligence-\(mail\)-assessment-will-shed-light-on-whether-young-students-have-had-opportunities-to-learn-and-to-engage-proactively-and-critically-in-a-world-where-production,-participation,-and-social-networking-are-increasingly-mediated-by-digital-and-ai-tools-/PISA%202029%20MAIL%20Assessment%20Framework%20First%20Draft.pdf](https://www.oecd.org/content/dam/oecd/en/about/projects/edu/the-pisa-2029-media-and-artificial-intelligence-(mail)-assessment-will-shed-light-on-whether-young-students-have-had-opportunities-to-learn-and-to-engage-proactively-and-critically-in-a-world-where-production,-participation,-and-social-networking-are-increasingly-mediated-by-digital-and-ai-tools-/PISA%202029%20MAIL%20Assessment%20Framework%20First%20Draft.pdf)
- Okubo, T. (2024). Towards More Diverse and Flexible International Large-Scale Assessments. *OECD Education Working Paper No. 310*.
- Plomp, T. J. (1998). *The potential of international comparative studies to monitor quality of education. Prospects*.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). *Advanced quantitative techniques in the social sciences: Vol. 1*. Thousand Oaks: Sage Publications.
- Robitaille, D. F., & Beaton, A. E. (2002). *Secondary Analysis of the TIMSS Data*. Dordrecht: Springer Netherlands.
- Ross, K. N., & Genevois, I. (2006). *Policy Forum on Cross-national Studies of the Quality of Education: Planning their Design and Managing their Impact*. Paris: International Institute for Educational Planning.
- Rutkowski, D., & Rutkowski, L. (2013). Measuring Socioeconomic Background in PISA: One Size Might not Fit all. *Research in Comparative and International Education*, 8(3), 259–278. <https://doi.org/10.2304/rcie.2013.8.3.259>
- Rutkowski, L., & Rutkowski, D. (2010). Getting it 'better': The importance of improving background questionnaires in international large-scale assessment. *Journal of curriculum studies*, 42(3), 411–430. <https://doi.org/10.1080/00220272.2010.487546>
- Sachs, G. D., Lafortune, G., Fuller, G., & Iablonovski, G. (2025). *Financing Sustainable Development to 2030 and Mid-Century: Sustainable Development Report 2025*. Dublin University Press Dublin, Ireland.

- SACMEQ (2023). *SACMEQ IV International Report*. Retrieved from http://www.seacmeq.org/sites/default/files/sacmeq/publications/sacmeq_iv_international_report.pdf
- Sahito, Z., & Väisänen, P. (2019). Teachers' Job Satisfaction in Developing Countries: A Literature Review. *European Journal of Cultural Studies*, 5(1).
- Sandoval-Hernández, A., Rutkowski, D., Tyler, M., & Daniel, M. (2019). Back to the drawing board: Can we compare socioeconomic background scales? *Revista de Educacion*. (383), 37–61.
- Scheerens, J. (2000). School effectiveness in developed and developing countries; a review of the research evidence. Retrieved from http://www.mp.gov.rs/resursi/dokumenti/dok20-eng-IIEP_school_effectiveness.pdf
- Scheerens, J. (1990). School Effectiveness Research and the Development of Process Indicators of School Functioning. *School Effectiveness and School Improvement*, 1(1), 61–80. <https://doi.org/10.1080/0924345900010106>
- Scheerens, J. (2004). Review of school and instructional effectiveness research. Background paper prepared for the Education for All Global Monitoring Report 2005: The Quality Imperative. Paris, UNESCO.
- Scheerens, J. (2016). Educational Effectiveness and Ineffectiveness. Advance online publication. <https://doi.org/10.1007/978-94-017-7459-8>
- Schmidt, W. H., & McKnight, C. C. (1995). Surveying Educational Opportunity in Mathematics and Science: An International Perspective. *Educational Evaluation and Policy Analysis*, 17(3), 337–353. <https://doi.org/10.3102/01623737017003337>
- Singh, R. (2021). *Teachers' Working Conditions in State and Non-State Schools*. Background Paper for the 2021/2 Global Education Monitoring Report.
- Smits, J., & Steendijk, R. (2015). The International Wealth Index (IWI). *Social Indicators Research*, 122(1), 65–85. <https://doi.org/10.1007/s11205-014-0683-x>
- Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *Kluwer international handbooks of education: Vol. 9. International handbook of educational evaluation* (pp. 31–62). Dordrecht, Boston: Kluwer Academic Publishers.
- Travers, K. J., & Westbury, I. (1989). *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*. Oxford: Pergamon Press.
- UIS (2025). *Analysis of Context Questionnaires in International Large-Scale Assessments (ILSAs)*. Retrieved from https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2025/02/GAML11_3.2_Context-questionnaires_ILSA.pdf
- UNESCO (2020). *Global Education Monitoring Report 2020: Inclusion and education: All means all*. Paris: UNESCO Publishing.
- UNESCO (2023). *Global Education Monitoring Report 2023: Technology in education: A tool on whose terms?*
- UNESCO (2024). *Assessments for Minimum Proficiency Levels a and b (AMPL-ab)*: UNESCO-UIS.
- UNESCO (2025). *Languages matter: Global guidance on multilingual education*.

- UNESCO, UNICEF, World Bank, OECD: GPE, & ACER (2021). *Contextual framework: COVID-19—Monitoring impacts on learning outcomes (MILO): Version 1.0*.
- UNICEF (2019). *SEA-PLM 2019 Assessment Framework*. Bangkok, Thailand.
- United Nations (2015). *Transforming our world: The 2030 Agenda for Sustainable Development* (A/RES/70/1). Retrieved from <https://undocs.org/en/A/RES/70/1>
- United Nations Children's Fund (2021). *The State of the World's Children 2021: On My Mind: Promoting, protecting and caring for children's mental health*. Retrieved from <https://www.unicef.org/media/114636/file/SOWC-2021-full-report-English.pdf>
- Van de Vijver, F. J. R., Jude, N., & Kuger, N. (2019). Challenges in International Large-Scale Educational Surveys. In L. E. Suter, E. Smith, & B. D. Denman (Eds.), *The SAGE Handbook of Comparative Studies in Education*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd.
- Wagemaker, H. (2020a). *Reliability and validity of international large-scale assessment: Understanding IEA's comparative studies of student achievement*.: Springer International Publishing.
- Wagemaker, H. (2020b). *Reliability and Validity of International Large-Scale Assessment*: Springer Nature.
- Wagemaker, H. (Ed.) (2020c). *Reliability and Validity of International Large-Scale Assessment: Understanding IEA's Comparative Studies of Student Achievement* (Vol. 10). Cham: Springer International Publishing.
- Wang, D., & Fawzi, W. W. (2020). Impacts of school feeding on educational and health outcomes of school-age children and adolescents in low- and middle-income countries: Protocol for a systematic review and meta-analysis. *Systematic Reviews*, 9(1), 148. <https://doi.org/10.1186/s13643-020-01317-6>
- Wartenberg, G., Aldrup, K., Grund, S., & Klusmann, U. (2023). Satisfied and High Performing? A Meta-Analysis and Systematic Review of the Correlates of Teachers' Job Satisfaction. *Educational Psychology Review*, 35(4), 179. <https://doi.org/10.1007/s10648-023-09831-4>
- Willms, J. D. (2018). *Learning divides: Using data to inform educational policy* (Information Paper No. 54). Retrieved from UNESCO Institute for Statistics website: <https://www.uis.unesco.org/sites/default/files/medias/fichiers/2025/08/ip54-learning-divides-using-data-inform-educational-policy.pdf>
- Willms, J., & Tramonte, L. (2015). Towards the development of contextual questionnaires for the PISA for development study. *OECD Education Working Papers*, No. 118. OECD Publishing. Retrieved from <https://doi.org/10.1787/5js1kv8crsjf-en>
- World Bank (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Retrieved from <https://www.worldbank.org/en/publication/wdr2018>
- World Bank (2025). *FY26 Updated Country Income Classification for Analytical Purposes (English)*. Retrieved from <http://documents.worldbank.org/curated/en/099062525111074500>
- World Bank Group (2026). *From Prospective to Prepared Teacher: A Global Study of Initial Teacher Education*.

Zhang, H., & Holden, S. T. (2023). *Country Socio-economic Development and Disparity in School Children's Reading Skills Learning in Africa*. Norwegian University of Life Science Centre for Land Tenure Studies.