

Chapter 2

Methodology for the review of international large-scale assessments in education

This chapter describes the methodology used for the review of the large-scale assessments. The chapter explains how the three main objectives of PISA for Development (PISA-D) were used as a framework for analysing the different surveys and extracting key principles, guidelines, approaches and lessons from the reviewed surveys' experiences, professional testing practices and approaches to provide recommendations and guidance for the PISA-D project.

Methodology

The review focused on the component skills assessed and contextual data collection used in relevant international assessments. Its methodology, like this resulting expert paper, has four main components:

1. A literature review of relevant literature on the international assessments, including reference materials specified in the terms of reference for this project, and the expert papers on technical strands:
 - a. cognitive instruments (Adams and Cresswell, 2014)
 - b. contextual questionnaires (Willms and Tramonte, 2014)
 - c. out-of-school 15-year-olds (Carr-Hill, 2015).
2. Personal communication with assessment agencies, to receive assessment material and information about procedures that are not publicly available (for example, test items and booklets, context questionnaires, details of the translation process). In accordance with the terms of reference, all communication was facilitated by the OECD and depended on the collaboration of the agencies and people addressed.
3. Analysing the review results and identifying the main findings according to the objectives of PISA-D.
4. Drafting options for PISA-D survey development and fieldwork based on the findings of the review.

The following three areas are a key part of the international assessments review:

1. Component skills and cognitive instruments, comprising: assessment frameworks (evolution and development, including main responsibility and participating countries' involvement, definitions of the domains), item development, test design (organisation, domain and framework coverage, item difficulty, test targeting and mode of delivery), psychometric analyses, scaling, calibration and equating methods, cross-country comparability, trends, proficiency levels, translation, adaptation and verification (including language of assessment, translated languages), field trial and item selection.
2. Contextual data collection instruments, comprising: evolution and development (including main responsibility, involvement of participating countries), content and question types (types of questionnaires, main theoretical constructs, variables, indices and scales, question formats, mode of delivery), questionnaire scales and technical aspects (including scale calibration methods and methods to ensure cross-cultural validity), translation, adaptation and verification process, field trial and item selection.
3. Implementation, methods and approaches to include out-of-school children, and use of data, comprising: implementation (institutional arrangements, survey operations and standardisation), methods and approaches for including out-of-school children (population and sampling, other considerations), and use of data (reports and communication of findings, use of results).

In addition to these key areas, the assessments' main characteristics were included in the review to get an understanding of each assessment programme's overall purpose and nature, but these will not be part of the analysis.

Survey classification

The international surveys were organised into three categories for the purpose of this report: large-scale international surveys, school-based surveys and household-based surveys.

The “large-scale international surveys” category covers international assessments and surveys that aim to produce internationally comparable datasets. To ensure international comparability, large-scale international surveys are highly standardised for all phases of the study, ranging from framework and instrument development, translation and verification procedures, test design, sample design, field operations, scaling methodology, data processing and management to quality assurance. PISA, which serves as a reference for this review, can be regarded as a typical example of a large-scale international survey. The surveys placed in this category for the review are therefore similar to PISA, especially in their level of standardisation. Large-scale international surveys included in this review are: PIRLS and prePIRLS, TIMSS and TIMSS Numeracy, SACMEQ, PASEC, LLECE and WEI-SPS. PrePIRLS and TIMSS Numeracy are specifically targeted at primary school children in developing countries and may provide valuable aspects for PISA-D.

Two school-based surveys were reviewed: EGRA and EGMA. Both surveys build on centrally developed expert frameworks, a toolkit and guidance notes for planning and implementation on the national level. The main aim is to measure the most basic literacy and mathematics skills in the early grades. Assessment tools are developed for each country context separately, and therefore EGRA and EGMA only allow limited comparison across countries. The surveys are carried out orally and in one-on-one settings in schools. In regard to their focus on the national level, EGRA and EGMA are similar to some of the surveys classed as household-based; however, their target population is students in schools and the tests are solely administered in schools.

The household-based surveys reviewed are PIAAC, STEP, LAMP, ASER and Uwezo. PIAAC, STEP and LAMP yield data that can be used on a national level as well as for international comparison – and therefore require a high level of standardisation, similar to international large-scale surveys. ASER and Uwezo focus on a national or district level. LAMP, ASER and Uwezo are administered to individuals in face-to-face interviews using mainly pencil and paper. The target population and sampling approaches of these surveys may contain valuable information for including the out-of-school population in PISA-D.

A summary of the main characteristics of the reviewed international surveys is given in Annex A.

PISA for Development participating countries

The countries that have agreed to participate in PISA-D as of November 2014 are Cambodia, Ecuador, Guatemala, Paraguay, Senegal and Zambia. Table 2.1 shows the current PISA-D countries by geographical location, status as official development assistance (ODA) recipients according to the OECD Development Assistance Committee (DAC), and participation in the international surveys reviewed.

Table 2.1 Countries participating in PISA-D according to geographical location, DAC-ODA recipient-status and participation in the international surveys reviewed

		Northwest South America	Western South America	Central America	Southeast Asia	West Africa	Southern Africa
		Ecuador	Paraguay	Guatemala	Cambodia	Senegal	Zambia
		Upper middle-income countries		Lower middle-income countries		Least developed countries	
Large-scale international surveys	PISA						
	PIRLS						
	PrePIRLS						
	TIMSS						
	SACMEQ						X
	PASEC				X	X	
	LLECE	X	X	X			
WEI-SPS		X					
School-based surveys	EGRA			X	X	X	X
	EGMA						X
Household-based surveys	PIAAC						
	STEP						
	LAMP		X				
	ASER						
	Uwezo						

Note: Regarding DAC-ODA recipient-status, effective reporting on 2012 and 2013 flows was valid until 31 December 2014. The position of these participating countries stays the same in the list that is effective as at 1 January 2015 for reporting on 2014, 2015 and 2016 flows.

Source: OECD, 2014a.

As shown in Table 2.1, the two African countries and Cambodia are among the least developed countries; Guatemala and Paraguay are among lower middle-income countries; and Ecuador among upper middle-income countries (classified by gross domestic product [GDP] per capita). According to Bloem (2013: 11), the majority of countries that have participated in PISA (all cycles up to 2015) are high and upper middle-income countries. There are eight lower middle-income countries and economies participating in PISA, and one low-income country (Kyrgyzstan). Participation from partner countries on the African continent has been limited so far to Algeria, Mauritius and Tunisia, which are among the upper middle-income countries. The number of partner countries participating in PISA in East and South Eastern Asia, Central Asia and Central and Eastern Europe, Latin America and the Caribbean has grown over the last two cycles. Middle Eastern countries and economies that have participated in PISA include Jordan, Qatar and the United Arab Emirates (Bloem, 2013: 8).

None of the PISA-D countries have participated in PISA before. All PISA-D countries have experience of at least one large-scale international survey: Ecuador and Guatemala participated in LLECE, Cambodia and Senegal in PASEC, and Zambia in SACMEQ. All countries apart from Ecuador have implemented the school-based survey EGRA, and Zambia also implemented EGMA.

The participation of developing countries in PISA presents a number of challenges to both the countries and PISA. Examples of these challenges concern funding, lack of institutional capacity and other issues regarding national implementation (due to the target population or the language of the test), performance coverage, relevance of performance results and “fear of bad performance”, lack of analytical capacity, use of results and diverging policy priorities, lack of capacity for a full international participation in international meetings, and training (Bloem, 2013: 18).

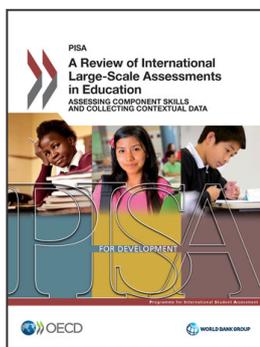
PISA-D looks for ways to overcome these challenges and to provide an assessment for developing countries that they can build on to improve their education system. PISA-D countries therefore require an assessment that (OECD, 2014b):

- Reports results on the PISA scale and supports comparability with international PISA results.
- Allows students to demonstrate the full range of proficiency levels.
- Provides policy-relevant information on students at the lower ends of proficiency levels.
- Adheres to PISA standards and identifies aspects that may be adjusted, while ensuring robustness of international comparability of results.

These requirements will strongly inform the following discussion of the review’s main findings for international assessments, and the options derived for PISA-D.

References

- Adams, R., and J. Cresswell (2014), *PISA for Development Technical Strand 2: Enhancement of PISA Cognitive Instruments*, OECD, Paris, [www.oecd.org/callsfortenders/Annex D - Cognitive Instruments.pdf](http://www.oecd.org/callsfortenders/Annex_D_-_Cognitive_Instruments.pdf).
- Bloem, S. (2013), “PISA in low and middle income countries”, *OECD Education Working Papers*, No. 93, OECD, Paris, <http://dx.doi.org/10.1787/5k41tm2gx2vd-en>.
- Carr-Hill, R. (2015), “PISA for Development Technical Strand C: Incorporating out-of-school 15- year-olds in the assessment”, *OECD Education Working Papers*, No. 120, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js0bsln9mg2-en>.
- OECD (2014a), “OECD list of ODA recipients”, www.oecd.org/dac/stats/daclistofodarecipients.htm (accessed 4 August 2014).
- OECD (2014b), *Call for Tender 100000990 - PISA for Development Strand A and Strand B*, OECD, Paris, <http://tinyurl.com/prjry24>.
- Willms, J.D. and L. Tramonte (2014), “Towards the development of contextual questionnaires for the PISA for development study”, *OECD Education Working Papers*, No. 118, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5js1kv8crsjf-en>.



From:

A Review of International Large-Scale Assessments in Education

Assessing Component Skills and Collecting Contextual Data

Access the complete publication at:

<https://doi.org/10.1787/9789264248373-en>

Please cite this chapter as:

Cresswell, John, Ursula Schwantner and Charlotte Waters (2015), "Methodology for the review of international large-scale assessments in education", in *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data*, The World Bank, Washington, D.C./OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264248373-5-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.