

Executive summary

The OECD has initiated PISA for Development (PISA-D) in response to the rising need of developing countries to collect data about the performance of their education systems and the capacity of their student bodies and in the context of the Education 2030 agenda which emphasises improved learning outcomes. This report has been commissioned by the OECD and the World Bank to inform the development and implementation of PISA-D but it also serves a wider interest in the experiences and lessons from the major international, regional and national large-scale educational assessments.

This report reviews the major large-scale learning assessments, including school-based surveys and household-based surveys. It aims to compare and contrast approaches regarding the instruments that are used to collect data on (a) component skills and cognitive instruments, (b) contextual frameworks, and (c) the implementation of the different international assessments, as well as approaches to include children who are not at school, and the ways in which data are used. It then seeks to identify assessment practices in these three areas that will be useful for the OECD and developing countries. For each of the issues discussed, there is a description of the prevailing international situation, followed by a consideration of the issue for developing countries and then a description of the relevance of the issue to PISA-D. A summary of the main characteristics of the reviewed surveys is given in Annex A.

The study makes many **recommendations**, particularly in respect of PISA-D, and the main ones are summarised as follows.

Component skills and cognitive assessment

- **Assessment frameworks:** For developing countries, it will be essential that any assessment has an agreed framework which has been arrived at through a process of discussion and negotiation, guided by experts in the field and by the countries participating in the assessment.
- **Item development:** Across the major international assessments there is a well-established procedure for the creation of new items. The procedure for item development in developing countries should follow this process.
- **Test design:** PISA-D should use a rotated booklet design allowing different students to be assessed on different parts of the framework.
- **Psychometric analyses, scaling, calibration and equating methods:** In developing countries, item response theory will deliver an accurate picture of student capacity across a wide range of item difficulties. It is recommended that the parameters used in scaling regular PISA should be adopted for PISA-D.

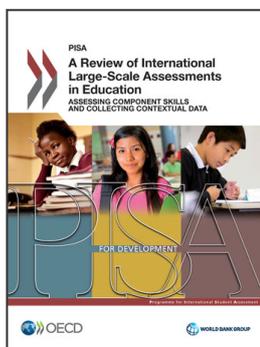
- **Cross-country comparability:** A differential item functioning process is usually undertaken at the field trial stage to identify any items that give an advantage or disadvantage to a particular country. It is recommended that PISA-D undertake a similar process.
- **Trends:** It will be important for developing countries to be able to quantify improvements by using assessments which include some of the same items from one test administration to the next.
- **Proficiency levels:** In PISA-D, an appropriately targeted test and the subsequent division of students into the various proficiency levels will provide extremely valuable information to the education ministries in the participating countries.
- **Translation, adaptation and verification of cognitive instruments:** It is recommended that the PISA-D project adopts the highest standards now operating in global assessments: that is, the double translation method.

Contextual data collection instruments

- **Types of contextual data collection instruments and mode of delivery:** PISA-D should give careful consideration to the types of questionnaires implemented, in order to collect the most essential contextual information in the most efficient way. It will be important to calculate a cost/value ratio for various contextual data collection instruments.
- **Translating, adapting and verifying contextual data collection instruments:** It is important to consider which languages are the most appropriate ones for the different groups of respondents. Questionnaires are preferably translated into the languages in which students, teachers, principals and parents are expected to be proficient.
- **Main factors and variables:** Regarding *early learning opportunities*, the PIRLS and TIMSS Learning to Read Survey (for parents), the LLECE questions about early reading and how often someone at home reads aloud to the child, and the questions about out-of-school status from ASER and Uwezo may all be of interest to PISA-D. Regarding *language at home and school*, a number of assessments contain items that may be relevant for PISA-D. For example, PIRLS and TIMSS contain questions about the frequency of speaking the language of the test at home and the language spoken by the student before school enrolment. PIRLS and TIMSS also ask if the books at home (“books at home” as used as an indicator for socio-economic status) are mainly in the test language.
- **Technical aspects of contextual data collection instruments:** Regarding question formats, PISA-D should include item formats that allow for an adjustment of self-reported measures. PISA-D should also undertake analyses to examine the extent of different patterns of response styles in participating countries.
- **Socio-economic status and poverty-related measures:** The surveys reviewed contain several good examples for socio-economic status (SES) and poverty-related measures relevant to PISA-D. SACMEQ, PASEC and LLECE include SES-related indices. SACMEQ and WEI-SPS include school and classroom measures that are related to SES.

Implementation procedures, methods and approaches to include out-of-school children, and use of data

- **Sampling:** Some countries do not maintain complete and up-to-date lists of schools. PISA-D will need to construct a school sampling frame that satisfies PISA's technical standards in these countries.
- **Data collection:** PISA-D should consider interview sessions to collect contextual data from respondents other than students. These respondents might include principals and teachers. It may be useful to implement: a tablet-based data collection tool to eliminate recording errors; cognitive test administration over multiple days; permitting extra time to complete cognitive assessments; establishing on-site test administrator checks of student booklets to reduce the incidence of missing/discrepant data; and sourcing test administrators who are local to the sites of test administration as a means of securing community engagement and buy-in.
- **Standardising implementation:** Articulation of standards could be included in memoranda of understanding or project implementation plans, as well as in a dedicated standards document. Including the standards in documents that are specific to each participating country, rather than general documents, may assist each country to be fully aware of its responsibilities with respect to the standards. A description of standards could be used as an opportunity to reflect the project's underlying values and ideology in a way that will help to secure local commitment to the project and acceptance of its results.
- **Methods and approaches to include out-of-school children:** Input should be sought from ASER and Uwezo and perhaps the other household-based assessments about how often they encounter problems with outdated sampling frames and how these are dealt with.
- **Analysis, reporting and use of data:** The use of benchmarks in the reviewed surveys should be examined. PISA-D should consider whether benchmarks might be incorporated into PISA-D analysis and reporting. Benchmarks that define minimum expected levels of performance may become increasingly relevant in the context of the post-2015 development goals and targets for education quality.



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), "Executive summary", 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/10.1787/9789264248373-3-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.