



# What have been the capacity-building outcomes for countries participating in PISA?

In order for countries to carry out more effective assessments of student learning, they must build or enhance country capacity in three areas: the enabling context for assessment, the alignment of all aspects of the education system, and assessment quality. The first section of this chapter examines how capacity has been measured, drawing on World Bank Systems Approach for Better Education Results (SABER)-Student Assessment reports. The second section explores how development partners and multilateral donors have helped build capacity for assessment, through training, technical publications and hands-on capacity building. The third section examines the evidence linking these activities to improved capacity, focusing on evidence from various international large-scale assessments and the Programme for International Student Assessment (PISA). This section also explains the capacity development activities of PISA for Development.

## WHAT HAVE BEEN THE CAPACITY-BUILDING OUTCOMES FOR COUNTRIES PARTICIPATING IN PISA?

## INTRODUCTION

In order for countries to carry out more effective assessments of student learning they must build capacity in three key areas: enabling context, system alignment and assessment quality (Clarke, 2012). The enabling context includes: setting clear policies for assessment, having strong public engagement for assessment, having regular funding for assessment, having strong organisational structures for assessment and having effective human resources for assessment. System alignment includes: aligning the assessment with system learning goals, curricula, textbooks and teacher development. Assessment quality includes: ensuring the technical quality of the assessment (from design to implementation to analysis and reporting) and effective use of the results.

Within each of these broad categories there are many specific capacity elements that involve building institutional, organisational and individual capacities. This chapter explores the extent to which participation in PISA has led to improved capacity in two of these broad areas: the enabling context and the technical quality of assessments. System alignment issues are more specific to a particular country and are best addressed in relation to national and sub-national assessments that are linked to a country's official learning goals and curricula.

## MEASURING ASSESSMENT CAPACITY

Systematic empirical measurements of a country's capacity for learning assessment are available for relatively few countries. More often, capacity is described in general terms. For example, one report reviews in detail the educational assessment systems of 22 Latin American countries with respect to 7 categories of analysis and 27 distinct themes, but provides no systematic summary of the results (Ferrer, 2006).

To obtain a systematic measurement of assessment capacity, The World Bank's Systems Approach for Better Education Results (SABER)-Student Assessment programme has assisted more than 20 countries to carry out reviews of their assessment capacity. These countries cover five regions (Europe and Central Asia, the Middle-East and North Africa, sub-Saharan Africa, South Asia and East Asia) and include three of the case study countries (Jordan, Kyrgyzstan, and Viet Nam). Similarly, the OECD and the World Bank have supported detailed capacity needs analyses for the six countries participating in the PISA for Development pilot programme, four of which have been completed (OECD, 2014a, 2014b, 2015a, 2015b). Both the SABER-Student Assessment and capacity needs analysis exercises evaluated countries on a specific set of capacities and then rated them as "latent", "emerging", "established" or "advanced" in each capacity area.

Results from these two systematic approaches show considerable need for capacity development in the area of large-scale assessments, with most of the ratings falling into the "emerging" category. Among the 20 countries with published SABER-Student Assessment evaluations, none received an "advanced" rating for either national or international large-scale assessments. A rating of "established" was given to 10% of countries for national large-scale assessments and to 20% of countries for international large-scale assessments. A rating of "emerging" was given to 70% of countries for national large-scale assessments and to 45% of countries for international large-scale assessments. Finally, a rating of "latent" was given to 20% of countries for national large-scale assessments and to 33% of countries for international large-scale assessments.


— Table 4.1 —

## National and international large-scale assessment capacity ratings

|                    | Jordan<br>(draft as of 2014) |             | Kyrgyzstan<br>(as of 2009) |          | Viet Nam<br>(as of 2009) |          |
|--------------------|------------------------------|-------------|----------------------------|----------|--------------------------|----------|
|                    | NLSA                         | ILSA        | NLSA                       | ILSA     | NLSA                     | ILSA     |
| Overall rating     | Emerging                     | Emerging    | Emerging                   | Emerging | Emerging                 | Latent   |
| Enabling context   | Emerging                     | Emerging    | Emerging                   | Emerging | Emerging                 | Latent   |
| System alignment   | Emerging                     | Established | Emerging                   | Emerging | Established              | Emerging |
| Assessment quality | Emerging                     | Emerging    | Emerging                   | Emerging | Established              | Latent   |

Notes: Ratings are based on the numerical average for the categories rated on each dimension, ranging from 1 = latent to 4 = advanced. National large-scale assessment (NLSA) and international large-scale assessment (ILSA).

Source: SABER-Student Assessment website and country reports [saber.worldbank.org/index.cfm](http://saber.worldbank.org/index.cfm).

StatLink  <http://dx.doi.org/10.1787/888933294048>



SABER-Student Assessment ratings for the three areas of enabling context, system alignment and assessment quality are generally similar within a country. For example, two of the case study countries for which SABER-Student Assessment evaluations are available were rated as having “emerging” capacity for both national and international large-scale assessments across most of the three broad areas (Table 4.1). However, differences in capacity for undertaking national and international large-scale assessments were found for all categories in Viet Nam, and for system alignment in Jordan. At the time of the SABER-Student Assessment evaluation, Viet Nam had not yet participated in either the Trends in International Mathematics and Science Study (TIMSS) or PISA.

## BUILDING CAPACITY FOR LARGE-SCALE ASSESSMENTS

Donors have provided support to low- and middle-income countries for assessment capacity building in three main ways: *i*) short-term and advanced degree level training; *ii*) technical publications on assessment; and *iii*) “hands-on” assessment capacity building. Most of this support has been targeted at raising awareness and improving the quality of assessments. In addition, support for secondary analysis of data is provided by most of the sponsoring organisations, including the OECD.

### Short-term and advanced degree training

Throughout the 1990s, the World Bank Institute offered courses designed to build or enhance a country’s capacity for carrying out learning assessments (Murphy et al., 1996). From 2002 to 2007, the World Bank sponsored regional workshops on national assessments, first in Asia between 2002 and 2004 (Bangkok, Colombo, Hanoi and Vientiane), and then in Africa between 2004 and 2006 (Accra, Addis Ababa, Cairo, Kampala, Maseru and Pretoria). In 2007, the Asian and African participants were brought together to share their experiences and receive additional training on item analysis (V. Greaney, personal communication, 11 July 2015).<sup>1</sup>

In 2014, the World Bank and the International Monetary Fund (IMF) sponsored a three-day workshop on national and international assessments for ministry of education officials from 20 countries in the Middle East and North Africa, and the United Nations Education, Science and Culture Organization (UNESCO) organised a four-day workshop on large-scale assessments for the Network on Education Quality Monitoring in Asia-Pacific. These workshops were intended to raise awareness about large-scale learning assessments and their utility for education policy.

Universities, professional associations and assessment institutions provide learning opportunities relevant to PISA participants through advanced degree training and technical short-term courses related to large-scale assessments. Universities provide individuals with the opportunity to deepen their knowledge regarding the technical issues in large-scale assessments through post-graduate work in statistics and psychometrics. Bilateral and multi-lateral donor agencies have often financed the participation of graduate students from low and middle-income countries in these programmes.

Professional associations, such as the American Educational Research Association, also offer short term courses related to the analysis of international large-scale assessments. These courses are typically offered concurrently with the associations’ annual meetings. Scholarships have been provided by these associations to participants from low-income and lower-middle-income countries.

Not-for profit institutions have provided fee-based courses on data analysis. The Australian Council for Education Research (ACER) has offered courses on analysing PISA data. The Educational Testing Service (ETS) has offered courses on analysing the results from the Programme for the International Assessment of Adult Competencies (PIAAC). The International Association for the Evaluation of Educational Achievement (IEA) has provided training in secondary analysis of TIMSS data (Wagemaker, 2014). Since 2007, the IEA-ETS Research Institute (IERI) Academy has offered 20 courses on topics such as assessment design, item-response theory, Bayesian models, and population modelling. It has also provided scholarships for participants from low-income and lower-middle-income countries.

### Technical publications on assessment

Publications raise awareness and provide technical information regarding learning assessments. During the 1990s, the World Bank supported publications designed to raise awareness of the importance of learning assessments, particularly national assessments of learning (Kellaghan and Greaney, 2001; Murphy et al., 1996). The World Bank’s SABER-Student Assessment programme is intended to raise awareness of various types of student assessments.

More recently, the World Bank has supported the preparation of a series of five books intended to build capacity for the multiple activities involved in undertaking, analysing and using national learning assessments (Box 4.1).

## WHAT HAVE BEEN THE CAPACITY-BUILDING OUTCOMES FOR COUNTRIES PARTICIPATING IN PISA?

Box 4.1 **Five World Bank books designed to build capacity for national assessments**

- *National Assessments of Educational Achievement, Volume 1: Assessing National Achievement Levels in Education* (Greaney and Kellaghan, 2008).
- *National Assessments of Educational Achievement, Volume 2: Developing Tests and Questionnaires for a National Assessment of Educational Achievement* (Anderson and Morgan, 2008).
- *National Assessments of Educational Achievement, Volume 3: Implementing a National Assessment of Educational Achievement* (Greaney and Kellaghan, 2012).
- *National Assessments of Educational Achievement, Volume 4: Analyzing Data from a National Assessment of Educational Achievement* (Shiel and Cartwright, 2015).
- *National Assessments of Educational Achievement, Volume 5: Using the Results of a National Assessment of Educational Achievement* (Kellaghan, Greaney, and Murray, 2009).

Publications by other not-for-profit organisations have also covered awareness and capacity building for assessment in middle-income countries. For example, the Partnership for Educational Revitalization in the Americas (PREAL) has published reviews of large-scale assessment programmes in Latin America (Ferrer, 2006). The IEA, in co-operation with ETS, sponsors an open-access journal, *Large-Scale Assessments in Education*, designed to “contribute to the science of large-scale assessments, help disseminate state-of-the-art information about empirical research using these databases and make the results available to policy makers and researchers around the world.” (*Large-scale Assessments in Education*, 2015). The American Psychological Association (APA) lists over 50 journals in the English language that focus on various technical aspects of non-clinical educational assessments (APA Division 5, 2011). PISA’s National Project Manager’s manual, and the technical reports prepared for each PISA cycle, can also be considered capacity building publications.

### Hands-on assessment capacity building

Direct assessment capacity building has been provided to low- and middle-income countries in two ways: through financial support from donors for participation in international large-scale assessments, and through the on-going technical support accompanying participation in large-scale assessments.

Donor support has enabled low- and middle-income countries to participate in regional and international assessments, with participation often including direct training activities (Lockheed, 2013). For example, the Association for the Development of Education in Africa promoted capacity development for assessment through its partnership with the Southern African Consortium for Measuring Educational Quality (SACMEQ). UNESCO, through the International Institute for Educational Planning (IIEP), supported the first two cycles of SACMEQ and, through the Latin-American Laboratory for the Assessment of the Quality of Education (LLECE), the first two cycles of the Latin American assessment: the Regional Comparative and Explanatory Study (ERCE). The United Nations Development Programme (UNDP) assisted five Arab states to participate in TIMSS 2003. The World Bank, through its Development Grant Facility, supported 18 low- and middle-income countries’ participation in TIMSS 1999, 24 low- and middle-income countries’ participation in the Progress in International Reading Literacy Study (PIRLS) 2001 and/or TIMSS 2003, 20 middle-income countries’ participation in PIRLS 2006 and/or TIMSS 2007, and also provided support to SACMEQ and the *Programme d’Analyse des Systèmes Educatif de CONFEMEN* (PASEC).

Participating countries typically receive hands-on training when participating in an international large-scale assessment. This type of training is often operational, but enables participants to build skills relevant to the assessment. For example, at the 2014 National Project Managers’ meeting for PISA 2015, held in Dublin, Ireland, specialised “hands-on” training and individual consultations related to use of the sampling software for the assessment were provided.

Donors have also supported countries to prepare national reports, and researchers to conduct secondary analysis of PISA data. For example, the Analysis for Development Group (GRADE) in Peru helped with the preparation of the national report for PISA 2012 in Peru.

### Support for secondary analysis of data

Three types of support for secondary analysis have been provided by donors and sponsors of large-scale international assessments: high quality data and data documentation, fellowships for data analysis and grants for conference



participation. Sponsors of international large-scale assessments facilitate secondary analysis through providing high quality data and data documentation for researchers. Technical reports on the assessments are prepared and published on the websites of the IEA, the OECD, the LLECE and SACMEQ, among others. Reports typically include the questionnaires, codebooks and analytic programmes (typically in Statistical Analysis Software – SAS – or Statistical Package for the Social Sciences – SPSS) to facilitate analysis. In addition, clean data sets are also made available. In the case of SACMEQ advance registration is required to use the data.

Fellowships for data analysis have also been provided. For example, in 2012 the OECD launched the Thomas J. Alexander fellowship programme to encourage researchers to use OECD education data to study issues of quality and equity. By 2015, 21 researchers – including fellows from Brazil, Chile, the People’s Republic of China (hereafter ‘China’), Poland, South Africa and Turkey – had been awarded fellowships to work alongside OECD analysts on issues of education quality and equity.

Grants for conference participation have been used to stimulate research on education quality in low- and middle-income countries, using data from international large-scale assessments. For example, in 2005, PREAL, in collaboration with the World Bank and the Global Development Network (GDN), provided grants for researchers from developing and transition countries to participate in a conference on educational research in these countries. Nearly 200 full papers were submitted, subjected to peer review, and critiqued by expert researchers.<sup>2</sup> The highest quality papers were subsequently published (Lockheed and Cueto, 2006).

## RESULTS FROM CAPACITY DEVELOPMENT ACTIVITIES

Few international large-scale assessment capacity development activities for low- or middle-income countries have been formally evaluated. Formal evaluations have been conducted for IEA and SACMEQ capacity development activities, but not for PISA. The SABER-Student Assessment evaluations cannot be used to determine whether or not PISA participation has built assessment capacity as they were only conducted once, typically after a country had already participated in an international large-scale assessment, including PISA, so any change in assessment capacity due to PISA participation is not observable.

Moreover, the overlap between participation in PISA and TIMSS makes it difficult to disentangle the effects of PISA from those of TIMSS.<sup>3</sup> However, countries that have participated in international large-scale assessments received higher capacity ratings. For example, the three countries – Kazakhstan, Serbia and the United Arab Emirates – that had participated in both TIMSS and PISA prior to the SABER-Student Assessment evaluation received an “established” rating (3.0) for international large-scale assessment capacity. Also, the average rating for international large-scale assessment capacity was higher (2.5) for the 8 countries that had participated in either or both of the two international large-scale assessments than it was for the 10 countries that had not participated (1.4). These associations between participation and capacity ratings suggest that international large-scale assessment participation may have enhanced these countries’ assessment capacity. Alternatively, countries with existing assessment capacity may have been more likely to participate in international large-scale assessments.<sup>4</sup>

## Capacity development through participation in TIMSS, PIRLS and SACMEQ

Independent evaluations of IEA and SACMEQ activities, required by donors, provide some empirical evidence of impact (Elley, 2002; Gilmore, 2005; Lockheed, 2010; Murimba, 2005). These evaluations, however, did not address all three dimensions of capacity (enabling context, system alignment, assessment quality) and generally focused on assessment quality. The evaluations often concluded that participating in an international large-scale assessment built technical and operational skills for assessment for the professionals who participate in the assessment. The evaluations are limited from a methodological perspective as they are only based on the participants’ self-reports of impact.

Nevertheless, participants report positive outcomes. For example, Elley (2002) reported that “In the course of the TIMSS-R project, training sessions were held for national research co-ordinators (NRCs) and members of their team in such technical matters as sampling, test development, coding, data processing and the like, and most NRCs claimed to benefit a great deal from these sessions.” Similarly, Gilmore, in assessing the impact of participation in another IEA study (TIMSS/PIRLS) reported that “NRCs were asked to rate the special or on-the-job training that they had received during the projects. Without exception, all NRCs reported that the training in sampling, test development, questionnaire development, data management, and quality control were ‘very helpful’ or ‘helpful’” (Gilmore, 2005). Another evaluation of an IEA programme found that the IEA training and capacity-building activities were viewed as having contributed to institutional and individual capacities related to assessment operations (Lockheed, 2010). Finally, Murimba (2005) noted that participation in SACMEQ’s capacity-building initiatives in sampling, data collection and data analysis were positively evaluated.



## WHAT HAVE BEEN THE CAPACITY-BUILDING OUTCOMES FOR COUNTRIES PARTICIPATING IN PISA?

### Capacity development through participation in PISA

PISA has the potential to build or enhance capacity for all dimensions of capacity development: enabling context, system alignment and assessment quality. Systematic evidence that participation specifically in PISA has improved a country's assessment capacity is limited for three reasons: *i*) no formal evaluation has been carried out; *ii*) opting into PISA can be based on a country's existing capacity; and *iii*) the focus of most PISA training has been on ensuring that the assessment was carried out in accordance with existing standards.

#### ***PISA's impact on assessment capacity has not yet been evaluated***

No formal evaluation of PISA's impact on assessment capacity has been carried out. While the OECD notes that it has helped countries successfully complete the assessment, it has not conducted or sponsored an empirical evaluation of PISA's impact on countries' assessment capacity, particularly for the middle-income participants. There is some anecdotal evidence that PISA has contributed to improvements in the technical quality of assessments.

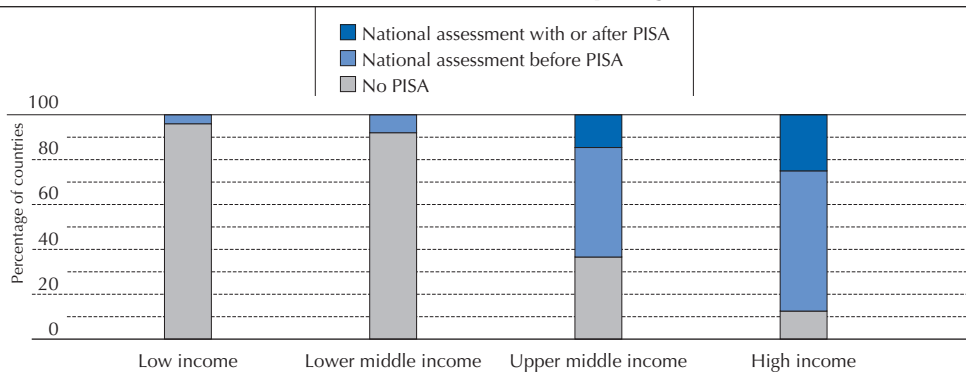
#### ***The OECD expects countries to have adequate assessment systems***

Initially, the OECD assumed that countries participating in PISA were capable of conducting the assessment according to OECD's standards and therefore did not measure the assessment capacity of the original participants. However, as more non-OECD countries applied to join PISA, the OECD has asked countries to demonstrate that they have the capacity to undertake the assessment, through, for example, the successful completion of a pilot assessment or field test in a PISA cycle. Participation in a prior international large-scale assessment, or evidence of a high quality national assessment, can also demonstrate capacity. The detail of PISA technical standards allows countries to self-select when applying for PISA participation, based on their understanding of these standards. Countries without sufficient existing individual or organisational technical capacity may choose not to apply.

Most of the countries that have participated in PISA have been high-income countries with relatively well-developed examination and assessment systems, and with considerable technical expertise in assessment. Participating countries also often have significant prior experience with other international and regional large-scale assessments, as well as national assessments, as demonstrated in Chapter 2.

— Figure 4.1 —

#### National assessments build capacity for PISA



Source: Author's analysis of data from Lockheed (2015), "Why do countries participate in international large-scale assessments? The case of PISA", *Policy Research Working Papers*, WPS 7447, World Bank, Washington DC.

StatLink <http://dx.doi.org/10.1787/888933293872>

Among the 145 countries that UNESCO (2015) identified as having national learning assessments, a high share of first-time PISA participants had several years of experience with national learning assessments. This was particularly the case for upper-middle- and high-income countries (Figure 4.1). About 70% of PISA participants, including 24 low- and middle-income countries, joined PISA after having conducted a national assessment. However, 16 upper-middle-income countries participated in PISA without having had prior national assessment experience.

#### ***PISA capacity development focused on technical quality***

The capacity development activities of prior PISA cycles have focused on ensuring that the assessment met the standards specified in the technical manuals. As a consequence, training focused on specific technical quality aspects of the assessment and was not designed to enhance the enabling context for assessment or directly contribute to system alignment.



Training has been closely embedded in the semi-annual meetings of the National Project Managers. Changes in the enabling contexts of participating countries may have occurred, however, and may be observed in the sector plans of some middle-income countries. For example, the sector plans of Cambodia refer to goals of performance on and participation in international assessments, and specifically mention PISA and international standards for learning. The evidence regarding PISA's impact on education policy, including agenda setting, is discussed in Chapter 5.

There is some evidence that the training activities related to PISA have built technical quality for assessment in countries. Interviews undertaken for this report with the PISA management teams from 16 low- and middle-income countries participating in PISA 2015 suggest that the PISA experience helped build capacity for conducting better-quality assessments through training related to specific operational tasks. The majority of the country teams, 12 out of 16, recalled that one or more member of the team had received training in selected areas. In response to the question: "At the time your country last participated in PISA, was your PISA unit or team offered opportunities to receive special or on-the-job training related to PISA?" the teams said that someone from the team had received training related to sampling (nine countries), data collection (nine countries), data entry (eight countries), professional scoring of constructed-response items (seven countries), and translation (seven countries).<sup>5</sup> Whether or not these training activities built assessment capacity is not known, but the responses indicate that training raised awareness about many of the technical elements of assessment.

While participation in PISA may have contributed to a country's assessment capacity development, no systematic evidence of this impact is available. Working with contractors and the OECD, countries opt into participating in PISA only after having determined that their assessment capacity is sufficient to undertake the assessment.

### **Building assessment capacity through PISA for Development**

The PISA for Development project includes specific activities designed to enhance assessment capacity, particularly the technical quality of the assessment. The OECD observes that the number of National Project Manager meetings has increased from six to eight, and their duration has increased from four to five days, compared to the main PISA NPM meetings. In addition, the final two meetings will be devoted to analysis topics and developing national reports. As with PISA, the meetings for PISA for Development will be held in participating countries, which facilitates the greater participation of operational staff rather than just the National Project Managers. Larger numbers of participants will also increase the likelihood of reaching a 'critical mass' of experienced staff that renders institutions more resilient to the effects of individual promotion, retirement and other migrations. Whether or not participation in the PISA for Development initiative will change the enabling context for assessment remains to be seen. A steering committee of stakeholders, often formalised as part of a PISA implementation plan, can extend the base of support for the assessment and may improve the enabling context in some countries. Because the capacity needs analysis includes indicators of the enabling context – setting clear policies for assessment, regular funding for the international large-scale assessment unit, and human resources – the impact of PISA participation on the enabling context for assessment can be addressed.

## **CONCLUSIONS AND IMPLICATIONS FOR PISA IN MIDDLE-INCOME COUNTRIES**

### **Conclusions**

Participants of international large-scale assessments perceive that participation improves a country's capacity for conducting high quality assessments. However, much remains to be learned about the impact of PISA participation on assessment capacity development in middle-income countries. In most participating countries, PISA implementation has benefitted from an existing assessment capacity that has been developed through participation in other assessments, particularly other international-large scale assessments. The experience of OECD countries suggests that participation in PISA raises awareness of technical standards and improves the quality of country-level PISA reports.

To date, however, there has been no rigorous evaluation of the impact of PISA participation on assessment capacity development in middle-income countries, and the available evidence is limited and mostly anecdotal. Some middle-income countries report that participation has built operational capacity for assessments and that they have introduced new techniques into their ongoing national assessments.

PISA for Development offers an opportunity to obtain empirical evidence regarding the effect of participation in PISA on assessment capacity, since participating countries undertake a capacity needs analysis prior to participation (Box 4.2). This allows countries to identify dimensions of assessment capacity that need strengthening; it also informs the OECD about areas needing enhanced training efforts. A follow-up capacity needs analysis could allow for a comparison of the country's assessment capacity before participating in PISA for Development with its capacity following participation.

## WHAT HAVE BEEN THE CAPACITY-BUILDING OUTCOMES FOR COUNTRIES PARTICIPATING IN PISA?

## Box 4.2 Evaluating PISA's capacity building impact

The PISA for Development initiative offers an opportunity to evaluate the impact of PISA participation on a country's assessment capacity. Baseline indicators of assessment capacity, which are crucial for assessing change, have been established through capacity needs analyses (CNAs) that have been completed in six of the seven pilot countries. The CNAs are structured around three levels of capacity: enabling context, organisation and individual. A software application has been developed that enables a systematic assessment of the country's capacity to carry out PISA for Development (Cartwright (2015). This CNA application produces 117 ratings of specific capacity needs and has been successfully used in Cambodia, Ecuador, Guatemala, Paraguay, Senegal and Zambia. To assess the enabling context capacity needs (48 rating categories), the CNA application uses the SABER-Student Assessment (Clarke, 2012) rating categories for more than 50% of the categories. To assess organisation capacity needs (41 ratings) and individual capacity needs (28 ratings) the CNA application developed rating categories drawn from specific project requirements. Only six SABER-Student Assessment ratings categories are used for organisational capacity and only one is used for individual capacity. The CNA for the pilot countries will serve as an empirical baseline, needed for assessing the degree to which PISA participation builds assessment capacity at any of these levels or categories.

Source: Modified from OECD (2014b), *Capacity Needs Analysis Report: Zambia*, PISA for Development, OECD Publishing, Paris [www.oecd.org/pisa/aboutpisa/NEW\\_Pisa%20for%20Development\\_Zambia\\_FINAL\\_revised.pdf](http://www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf).

## Implications

Two main implications of this chapter are that: *i*) greater attention will need to be placed on capacity building activities for PISA; and *ii*) greater effort must be made to empirically evaluate the effects of PISA participation on assessment capacity.

### For middle-income countries

To understand the strengths and weaknesses of their own capacity for undertaking international large-scale assessments:

1. Countries should undertake systematic analyses of their assessment capacity. Many countries may not have a good understanding of the scope of assessment operations or of the detailed project management involved. An online tool, such as the CNA, may help countries not only identify the tasks involved in conducting an assessment, but also facilitate the management of these tasks.

### For development partners

To enhance capacity building activities:

1. Donors could focus on remediating the assessment weaknesses identified in the CNA's of low- and middle-income countries that are considering participating in PISA. Support (including resources and training) could target these weaknesses as they are identified for each participating country.
2. Donors could also consider supporting or enhancing online self-help tools mapped to the National Project Manager's manual, which could improve the efficiency of management and budgeting related to assessment. For example, the CNA includes a management tool that allows countries to: *i*) co-ordinate their in-country tasks with the international timeline; *ii*) add local tasks; *iii*) allocate resources to and identify dependencies among tasks; and *iv*) incorporate changes in the international timeline into country timelines. This type of tool could be valuable in communicating the complexity of the assessment process itself.

To provide empirical evidence of PISA's effect of assessment capacity:

1. Donors could commission empirical evaluation of the impact of PISA participation on assessment capacity development in middle-income countries. The preparation and design of the PISA for Development initiative offer opportunities to evaluate the impact of participating in PISA on a country's improvement in assessment capacity. Currently, baseline indicators of assessment capacity, which are crucial for assessing change in capacity, have been established through CNAs that have been completed in six of the seven PISA for Development pilot countries (OECD, 2014a, 2014b, 2015a, 2015b, 2015c, 2015 forthcoming). By administering the CNAs after the PISA cycle has been completed it will be possible to identify areas in which participation in PISA has been associated with change and improvement in capacity.<sup>6</sup>





## Notes

1. These workshops eventually led to the publications listed in Box 4.1.
2. Esther Duflo, Erik Hanushek, Anil Kanjee, Peter Orazem, Erlinda Pefianco and Paulo Renato Souza.
3. Prior to the SABER-Student Assessment evaluation, four countries had participated in TIMSS only, three countries had participated in PISA only, and three countries had participated in both of these international large-scale assessments.
4. SABER-Student Assessment evaluations for both Lebanon and Viet Nam were completed before they participated in PISA, so these two countries are omitted from this analysis.
5. The fact that around half of the respondent teams reported that their team had not received training is surprising, since all country teams should have received training.
6. A pre-post comparison of assessment capacity can be made, but would not meet standards for impact evaluations.

## References

- Anderson, P.** and **G. Morgan** (2008), *National Assessments of Educational Achievement, Volume 2: Developing Tests and Questionnaires for a National Assessment of Educational Achievement*, The World Bank, Washington DC.
- APA Division 5** (2011), Journals in Assessment, Evaluation, Measurement, Psychometrics and Statistics, [www.apa.org/divisions/div5/journals.html](http://www.apa.org/divisions/div5/journals.html).
- Cartwright, F.** (2015), *PISA for Development Planner*, <http://polymetrika.org/PISAD>.
- Clarke, M.** (2012), *What Matters Most for Student Assessment Systems: A Framework Paper*, The World Bank, Washington, DC.
- Elley, W.** (2002), *Evaluating the Impact of TIMSS-R (1999) in Low- and Middle-Income Countries: An Independent Report on the Value of World Bank Support for an International Survey of Achievement in Mathematics and Science*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.
- Ferrer, G.** (2006), *Educational Assessment Systems in Latin America: Current Practice and Future Challenges*, PREAL, Washington, DC.
- Gilmore, A.** (2005), *The Impact of PIRLS (2001) and TIMSS (2003) in Low- and Middle-income Countries*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.
- Greaney, V.** and **T. Kellaghan** (2012), *National Assessments of Educational Achievement, Volume 3: Implementing a National Assessment of Educational Achievement*, The World Bank, Washington, DC.
- Greaney, V.** and **T. Kellaghan** (2008), *National Assessments of Educational Achievement, Volume 1: Assessing National Achievement Levels in Education*, World Bank Publications, Washington DC.
- Kellaghan, T.** and **V. Greaney** (2001), *Using Assessment to Improve the Quality of Education*, International Institute for Educational Planning, UNESCO, Paris.
- Kellaghan, T., V. Greaney** and **S. Murray** (2009), *National Assessments of Educational Achievement, Volume 5: Using the Results of a National Assessment of Educational Achievement*, World Bank Publications, Washington DC.
- Large-scale Assessments in Education** (2015), [www.largescaleassessmentsineducation.com/about](http://www.largescaleassessmentsineducation.com/about).
- Lockheed, M.** and **S. Cueto** (eds.) (2006), "Education equity and quality in developing and transition countries", [Special Issue] *International Journal of Educational Research*, Vol. 45, No. 3. Elsevier, Amsterdam.
- Lockheed, M.** (2015), "Why do countries participate in international large-scale assessments? The case of PISA", *Policy Research Working Papers*, WPS 7447, World Bank, Washington DC.
- Lockheed, M.** (2013), "Causes and consequences of international assessments in developing countries", in H.-D. Meyer and A. Benavot (eds.), *PISA, Power and Policy: The Emergence of Global Educational Governance*, Symposium Books, Didcot, Oxford.
- Lockheed, M.** (2010), *The Craft of Education Assessment: Does Participating in International and Regional Assessments Build Assessment Capacity in Developing Countries?* International Association for the Evaluation of Educational Achievement (IEA), Amsterdam.
- Murimba, S.** (2005), "The Impact of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ)", *Prospects*, Vol. 35, No. 1, pp. 91-108. Springer.
- Murphy, P., V. Greaney, M. Lockheed** and **C. Rohas** (eds.) (1996), *National Assessments: Testing the System*, The World Bank, Washington DC.

**WHAT HAVE BEEN THE CAPACITY-BUILDING OUTCOMES FOR COUNTRIES PARTICIPATING IN PISA?**

OECD (2015a), *Capacity Needs Analysis Report: Guatemala*, PISA for Development, OECD, Paris, [www.oecd.org/pisa/aboutpisa/NEW\\_Guatemala%20CNA%20reportFINAL\\_SL.pdf](http://www.oecd.org/pisa/aboutpisa/NEW_Guatemala%20CNA%20reportFINAL_SL.pdf).

OECD (2015b), *Capacity Needs Analysis Report: Ecuador*, PISA for Development, OECD, Paris, [www.oecd.org/pisa/aboutpisa/NEW\\_Ecuador%20CNA%20ReportFINAL\\_SL.pdf](http://www.oecd.org/pisa/aboutpisa/NEW_Ecuador%20CNA%20ReportFINAL_SL.pdf).

OECD (2015c), *Capacity Needs Analysis Report: Paraguay*, PISA for Development, OECD, Paris, [www.oecd.org/pisa/aboutpisa/Paraguay%20CNA%20report\\_FINAL.pdf](http://www.oecd.org/pisa/aboutpisa/Paraguay%20CNA%20report_FINAL.pdf).

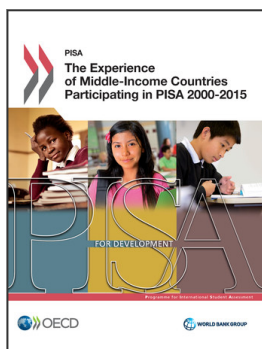
OECD (forthcoming), *Capacity Needs Analysis Report: Cambodia*, PISA for Development, OECD, Paris.

OECD (2014a), *Capacity Needs Analysis Report: Senegal*, PISA for Development, OECD, Paris, [www.oecd.org/pisa/aboutpisa/NEW\\_Pisa-D\\_CNA\\_Senegal\\_FINAL\\_revised.pdf](http://www.oecd.org/pisa/aboutpisa/NEW_Pisa-D_CNA_Senegal_FINAL_revised.pdf).

OECD (2014b), *Capacity Needs Analysis Report: Zambia*, PISA for Development, OECD, Paris, [www.oecd.org/pisa/aboutpisa/NEW\\_Pisa%20for%20Development\\_Zambia\\_FINAL\\_revised.pdf](http://www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf).

Shiel, G. and F. Cartwright (2015), *National Assessments of Educational Achievement, Volume 4: Analyzing Data from a National Assessment of Educational Achievement*, World Bank Publications, Washington, DC.

Wagemaker, H. (2014), "International large-scale assessments: From research to policy", in L. Rutkowski, M.V. Davier, and D. Rutkowski. (eds.), *Handbook of International Large-Scale Assessment: Background, Technical Issues and Methods of Data Analysis*, Taylor and Francis, New York.



**From:**  
**The Experience of Middle-Income Countries  
Participating in PISA 2000-2015**

**Access the complete publication at:**  
<https://doi.org/10.1787/9789264246195-en>

**Please cite this chapter as:**

Lockheed, Marlaine, Tijana Prokic-Bruer and Anna Shadrova (2015), "What have been the capacity-building outcomes for countries participating in PISA?", in *The Experience of Middle-Income Countries Participating in PISA 2000-2015*, The World Bank, Washington, D.C./OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264246195-7-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](mailto: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](mailto:info@copyright.com) or the Centre français d'exploitation du droit de copie (CFC) at [contact@cfcopies.com](mailto:contact@cfcopies.com).