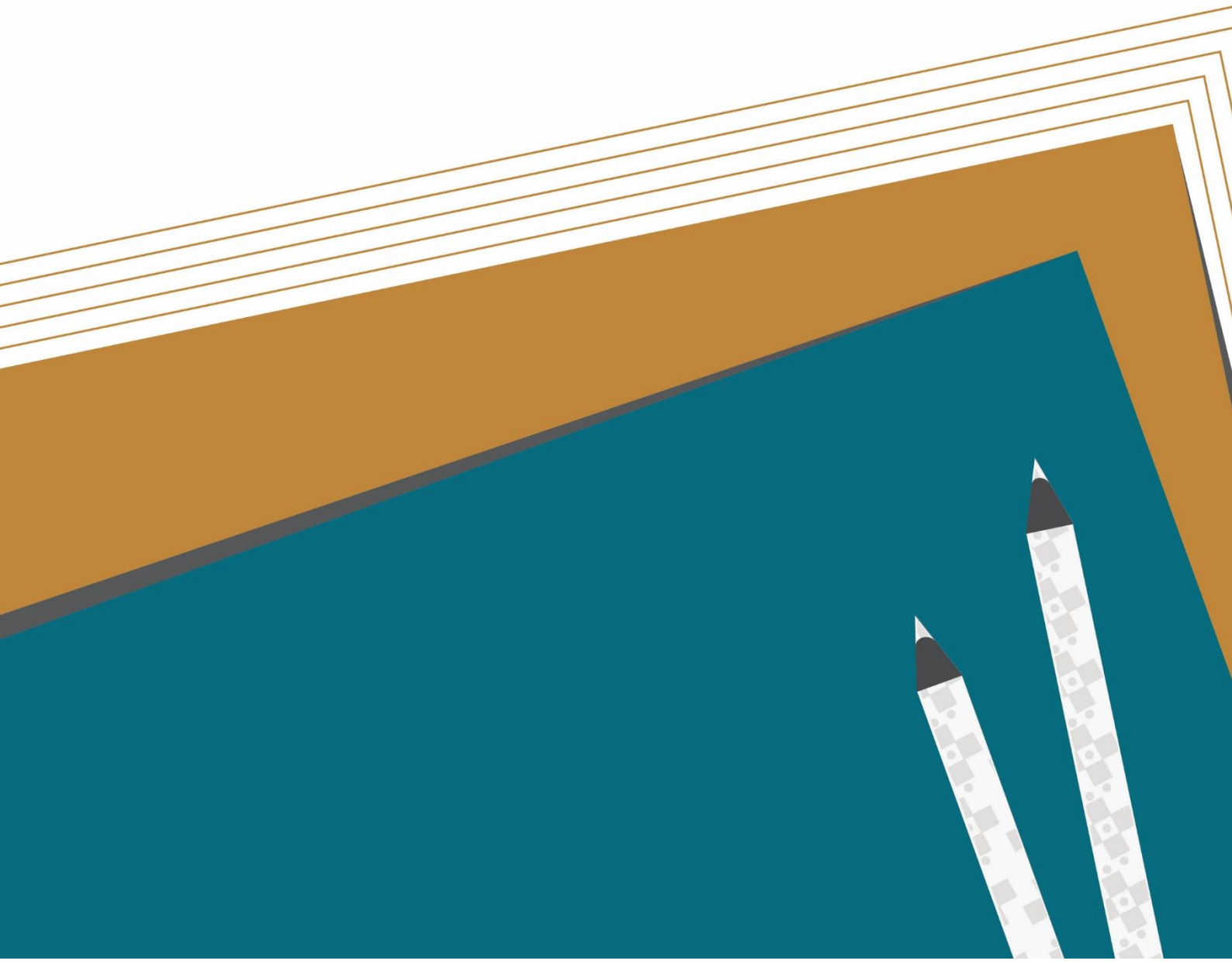
Azim Premji University

Using large-scale assessments to inform education policy and governance



Using large-scale assessments to inform education policy and governance

Abbreviations

ASER	Annual Status of Education Report
FPE	Free primary education
IEA	International Association for the Evaluation of Educational Achievement
IIEP	International Institute for Educational Planning
KESSP	Kenya Education Sector Support Programme
LSA	Large-scale assessment
NAS	National Achievement Survey
OECD	Organization for Economic Cooperation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
TIMSS	Trends in International Mathematics and Science Study
SAS	State Achievement Survey
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality

Introduction

Large-scale assessments (LSAs) are educational assessments that provide information about overall academic achievement in education systems. Most LSAs are sample based, where a representative student sample's performance is used to make inferences about the entire student population. Some LSAs are census based as they test the entire student population. LSAs may be subnational (SAS in Karnataka), national (ASER and NAS in India), regional (SACMEQ in Africa), or international in scope (PISA, TIMMS, and PIRLS). LSAs are always low-stakes assessments as the results do not affect the academic or career prospects of students.

LSAs may either test the school curriculum or particular content areas for a given age or grade level. For instance, PIRLS tests the reading achievement of fourth-graders. PISA tests the work-readiness of 15-year-olds based on their performance in reading, mathematics and science. In India, NAS tests the academic achievement of children in grades 3, 5, 8 and 10. Most LSAs also collect information on background variables that influence student performance. These background variables can include gender, income, teacher training and qualifications, infrastructure, region, school characteristics, etc.

The results of LSAs have many uses. Relating background variables to student performance can provide insights that help guide education policy and administration. Reforms based on LSAs often target quality, equity, accountability and resource allocation. The following vignettes illustrate how the results of LSAs can be used fruitfully.

Improving access to resources

Kenya participated in SACMEQ I and SACMEQ II in 1998 and 2000 respectively. SACMEQ I tested the reading achievement of grade 6 students, while SACMEQ II tested the reading and mathematics achievement of both students and teachers. In Kenya, SACMEQ results revealed a relationship between inputs (pens, pencils, notebooks and textbooks) and reading performance. The poorly resourced Western and North-Eastern provinces had the lowest reading levels. In contrast, well-resourced Nairobi had the best reading levels (Nzomo & Makuwa, 2006). Other findings of concern included relatively high rates of absenteeism, low female teacher representation and inadequate toilet facilities in poorer provinces, especially the North-Eastern provinces (Nzomo & Makuwa, 2006). The reasons were attributed to social inequality and a lack of well-defined norms for resource inputs. There were also disparities between official government standards and on-the-ground practices of resource allocation. To reduce disparities, the government developed a set of resource benchmarks for all schools (Nzomo & Makuwa, 2006).

The results of SACMEQ I and II influenced policy research documents as well. The recommendations of the documents informed programmes such as FPE in 2003 and KESSP in 2005 (Wasanga, Ogle and Wambua, 2012). The FPE programme aims at providing free primary education in Kenya. The provisions of FPE that target equity include the abolishment of tuition fees, relaxed age requirements, provision of teaching-learning materials and government funds for the most vulnerable children (Wasanga, Ogle and Wambua, 2012).

The KESSP aimed to improve the quality and accessibility of learning environments in schools (Wasanga, Ogle and Wambua, 2012). It involved collaborations between the government and communities, parents, NGOs and development partners. The KESSP involved targeted interventions and investments in areas such as health and nutrition, infrastructure, instructional materials, gender equality, special education, and non-formal education (Wasanga, Ogle and Wambua, 2012).

Improving educational equity

Germany experienced a "PISA shock" when the results of the first PISA cycle were released in 2000. Germany scored well below the OECD average for reading science and mathematics. There were also significant disparities in scores based on socio-economic and migration status (Davoli & Horst, 2018). Germany's disappointing performance in PISA prompted a wholesale reform of the education system.

Earlier, based on academic performance, students were streamed into lower secondary schools called *Gymnasia*, *Realschulen* and *Hauptschulen*[1] at the early age of 10. PISA 2000 demonstrated that streaming did not reliably segregate students based on ability because there were significant overlaps between the performance of students in *Gymnasia*, *Realschulen* and *Hauptschulen* (Kotthoff, 2011). When students with the same cognitive abilities were compared, it showed that students of an advantaged socio-economic background were more likely to attend *Gymnasia* than their peers (Kotthoff, 2011). More children in disadvantaged socio-economic groups (almost 40%) had level-1 reading skills compared to the most advantaged socio-economic groups (only 10%) (Kotthoff, 2011).

One of the most important reforms was the reduction of segregation in the German school system. Streaming was not completely abolished, but *Hauptschulen* schools were merged with *Realschulen* in many states. This reform decreased the stigma associated with *Hauptschulen*, which were attended by the most disadvantaged students (Davoli & Horst, 2018). Other reforms such as lengthening school days and expanding preschool attendance, too, were introduced (Davoli & Horst, 2018). Access to preschools has helped reduce the gap between children with more and less educated parents (Davoli & Horst, 2018). Migrant children started receiving German language training in preschool and kindergarten to address their challenges (Davoli & Horst, 2018).

Improving teacher quality

Peru experienced a "PISA shock" when it ranked last out of the 65 participating countries in the 2012 PISA cycle. The disappointing result prompted educational reforms in the areas of teacher professional development, improvement in educational quality, effective management of the school system and improving educational infrastructure (Saavedra & Gutierrez, 2020). Peru passed a Teacher's Reform Law in 2012 with the hopes of attracting quality candidates to the teaching

[1] *Gymnasia* are the most elite secondary schools of Germany that prepare students for university and academic credentials. *Realschulen* and *Hauptschulen* follow *Gynmasia* in terms of status, with *Realschulen* being considered better than *Hauptschulen*. Both *Hauptschulen* and *Realschulen* cater to students seeking a vocational occupation. However, unlike *Hauptschulen*, *Realschulen* provide foreign language certifications, access to higher vocational schools, and opportunities to transition to Gymnasia.

profession. In 2014, Peru offered full merit-based scholarships for undergraduates to study pedagogy in the best universities (Saavedra & Gutierrez, 2020). Reforms in teacher promotion were also instituted. Promotions were not defined by years of service but by performance in transparent, objective and fair exams. Teachers who passed promotion exams received relatively large bonuses (Saavedra & Gutierrez, 2020).

Investments were also made in teacher professional development. In 2016, the Peruvian government introduced a Teacher Induction Programme for those with less than two years of experience in public school teaching. Teachers are paired with experienced mentors for six months and given access to online training materials and remote guidance (Saavedra & Gutierrez, 2020).

Legitimising reform and improving accountability

Shanghai performed impressively in the 2009 and 2012 cycles of PISA. However, government officials were concerned about the long homework hours of students as reported in PISA assessments. Their concerns also stemmed from news reports of stress, burnout and health issues among students. The Chinese government officials used the results of PISA 2012 to legitimise reforms in quality appraisal (Tan, 2019).

In 2010, the Shanghai government introduced a system of quality appraisal called Green Indices for the Academic Quality of Primary and Secondary Schools. The Green Indices do not just measure academic achievement but also focus on students' moral conduct and physical and mental health (Tan, 2019). The indicators of the Green Indices are academic performance, learning motivation, schoolwork burden, teacher-student relationship, teaching methods, principal's leadership, students' moral conduct and physical and mental health (Tan, 2019).

The questionnaires used for the Green Indices evaluations are inspired by PISA and include questions about lifestyle factors such as hours of sleep and homework (Tan, 2019). However, unlike PISA, the Green Indices cover grades 4 and 9 along with subjects such as moral education, physical education and art. To ensure that the Green Indices are taken seriously, they are a part of school inspections, performance appraisals and five-year developmental plans.

Improving teaching and learning in schools

The Czech Republic analysed the results of TIMSS 2007 to identify the weakest concepts among its grade-4 students (Mullis & Martin, 2012). The analysis revealed that students lagged behind their peers from other countries in fractions and decimals. Items that posed particular difficulty to Czech students were also identified. Teachers' manuals were developed based on the common misconceptions and errors identified in the analysis (Mullis & Martin, 2012). These manuals included activities and tasks that would help teachers identify and remediate misconceptions

Reforming school curricula and assessments

Botswana participated in two cycles of TIMSS in 2003 and 2007. Its students underperformed in both cycles, and about 68% of them failed to reach the Low International Benchmark. In response, the government reviewed the upper primary and junior secondary curricula (Mullis & Martin, 2012). The revised curricula were benchmarked using the TIMSS mathematics and science frameworks. As the results of both TIMSS cycles showed poor performance in critical thinking and problem solving, emphasis on these skills in the revised curricula was increased (Mullis & Martin, 2012). Assessment objectives were also revised to include Higher Order Thinking Skills (HOTS) in the Primary School Leaving Examination and the Junior Certificate Examination.

Conclusion

Policymakers, administrators, teachers and educators can use the results of LSAs in many ways to improve education systems. LSAs help establish relationships between academic performance and influential factors such as income, gender, access to resources, etc. Identifying disparities can lead to reform and improved resource allocation. Deficiencies in teacher quality, school curricula, teaching-learning processes and assessments can be identified and corrected. Most importantly, making comparisons between countries can provide fresh perspectives on problems, necessary reforms and improvements for education systems in a given country or region.

References

Cresswell, J., Schwanter, U., & Waters, C. (2015). Annex A: General information about the international surveys reviewed. In *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data* (pp. 151-171). PISA, The World Bank, Washington D.C./OECD Publishing/Paris.

Davoli, M., & H. E. (2018, August). The PISA Shock, Socioeconomic Inequality, and School Reforms in Germany. *IZA Policy Paper No. 140*. Institute of Labour Economics (IZA).

Kotthoff, H.G. (2011). Between Excellence and Equity: The Case of the German Education System. *Revista Española de Educación Comparada, 18*, 27-60.

Mullis, I.V., & Martin, M.O. (2012). Using TIMSS and PIRLS to improve teaching and learning. *Recherches en éducation, 14*. doi:https://doi.org/10.4000/ree.5835

Nzomo, J., & Makuwa, D. (2006). How can countries move from cross-national research results to dissemination, and then to policy reform? In K.N. Ross, & I. Genevois (Eds.), *Cross-National Studies of the Quality of Education: Planning their Design and Managing their Impact* (pp. 213-225). Paris: International Institute for Educational Planning.

Saavedra, J., & Gutierrez, M. (2020). Peru: A Wholesale Reform Fueled by an Obsession with Learning and Equity. In F. M. Reimers (Ed.), *Audacious Education Purposes: How Governments Transform the Goals of Education Systems* (pp. 153-180). Cham, Switzerland: Springer Open.

Tan, C. (2019). PISA and education reform in Shanghai. *Critical Studies in Education, 60*(13), 391-406.

Wasanga, P.M., Ogle, M.A., & Wambua, R.M. (2012). *The SACMEQ III Project in Kenya: A Study of the Conditions of Schooling and the Quality of Education.* Nairobi: Kenya National Examinations Council.

Appendix

General Information about Some Popular Large-Scale Assessments (Cresswell, Schwanter, & Waters, 2015)

Name	Conducted by	Years conducted	Sample type	Countries	Target population	Content areas tested
International						
PISA	OECD	2000, 2003, 2006, 2009, 2012, 2015	Sample-based	65, including all OECD countries	15-year-olds	Compulsory: reading, mathematics, science. Optional: collaborative problem-solving
PIRLS	IEA	2001, 2006, 2011, 2016	Sample-based	49 in 2011	Grade 4	Literacy
TIMSS	IEA	1995, 1999, 2003, 2007, 2011	Sample-based	77 in 2011	Grades 4 and 8	Mathematics, science
Regional						

SACMEQ	SACMEQ, UNESCO-IIEP	1999, 2004, 2011, 2014	Sample-based	15 African countries	Grade 6	Literacy, numeracy, health
National						
ASER	Pratham	Annually since 2005	Sample-based	Pan-India	Children between ages 5 and 16	Reading, arithmetic
NAS	NCERT	Thrice-yearly from 2001 onwards	Sample-based	Pan-India	Grades 3, 5, 8 and 10	School subjects

This article is authored by Madhumati Manjunath. She is an intern at the School of Continuing Education and University Resource Centre (SCE-URC) at Azim Premji University. She can be reached at madhumati.manjunath19_mae@apu.edu.in.
This article can be cited as-

Using large-scale assessments to inform education policy and governance, Assessment resources, 2021, Azim Premji University

Using large-scale assessments to inform education policy and governance