Disseminating and Using Student Assessment Information in Jordan

Osamha Obeidat and Zaina Dawani
Disseminating and Using Student Assessment Information in Jordan

Osamha Obeidat and Zaina Dawani
Contents
Abbreviations and Acronyms .........................................................................................v
About the Series .............................................................................................................. vii
About the Authors ........................................................................................................... ix
Acknowledgments ............................................................................................................ xi
Abstract .......................................................................................................................... xiii
Executive Summary ......................................................................................................... xv
Introduction ....................................................................................................................... 1
1. Overview of the Student Assessment System ............................................................. 3
2. Classroom Assessment ................................................................................................. 11
3. General Secondary Examination (Tawjihi) ................................................................. 14
4. Large-Scale National Assessments ............................................................................ 17
5. International Assessments .......................................................................................... 26
6. Drivers for Change ....................................................................................................... 29
7. Discussion ...................................................................................................................... 31
8. Lessons Learned .......................................................................................................... 35
Conclusion ........................................................................................................................ 38
References ......................................................................................................................... 39

Figures
Figure 1. Assessment Institutions and Entities in Jordan................................................. 9
Figure 2. Evolution of Student Assessment System in Jordan ......................................... 11
Figure 3. National Test Report for Grade 8, 2012–2013 .................................................. 20
Figure 4. Example of Schools Report, 2012–2013 ............................................................ 21
Figure 5. NAfKE Policy Brief, 2013 ................................................................................ 24
Figure 6. NAfKE Results, 2011: Achievement Levels by Subject and Grade ............... 25
Figure 7. Summary of Assessment Programs in Jordan .................................................... 32
Figure 8. Percentage of MoE Staff who Read Evaluation Reports, Executive
Summary, and/or Policy Brief ......................................................................................... 33
Figure 9. Percentage of Field Directorate Staff who Read Evaluation Reports,
Executive Summary, and/or Policy Brief ....................................................................... 33
Figure 10. Were Findings Presented in Reports Useful in the Education
Decision-Making Process? ............................................................................................... 34

Tables
Table 1. Example of ERfKE M&E Framework Indicators ............................................... 5
Table 2. Achievement-Level Categories on the National Test ......................................... 18

Boxes
Box 1. Dissemination and Use of Classroom Assessment Information ....................... 12
Box 2. Methods of “Authentic Assessment” ................................................................... 13
Box 3. Dissemination and Use of Tawjihi Data and Results ....................................... 15

Disseminating and Using Student Assessment Information in Jordan
Box 4. Dissemination and Use of Data and Results from the National Test for the Control of Education Quality.................................................................................. 19
Box 5. Dissemination and Use of Data and Results from the NAfKE .................. 22
Box 6. Factors Contributing to Decline in Student Achievement on NAfKE .... 26
Box 7. Dissemination and Use of International Assessment Data and Results..... 27
**Abbreviations and Acronyms**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCU</td>
<td>Development Coordination Unit, Ministry of Education</td>
</tr>
<tr>
<td>DRP</td>
<td>Directorate of Research and Planning</td>
</tr>
<tr>
<td>EGMA</td>
<td>Early Grade Mathematics Assessment</td>
</tr>
<tr>
<td>EGRA</td>
<td>Early Grade Reading Assessment</td>
</tr>
<tr>
<td>EMIS</td>
<td>education management information system</td>
</tr>
<tr>
<td>EMISD</td>
<td>Education Management Information System Directorate, Ministry of Education</td>
</tr>
<tr>
<td>ERfKE</td>
<td>Education Reform for Knowledge Economy</td>
</tr>
<tr>
<td>ERSD</td>
<td>Education Research and Studies Division</td>
</tr>
<tr>
<td>ERSDir</td>
<td>Education Research and Studies Directorate</td>
</tr>
<tr>
<td>IEAP-II</td>
<td>International Assessment of Educational Progress II</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>monitoring and evaluation</td>
</tr>
<tr>
<td>MED</td>
<td>M&amp;E Division</td>
</tr>
<tr>
<td>MoE</td>
<td>Ministry of Education</td>
</tr>
<tr>
<td>NAfKE</td>
<td>National Assessment for Knowledge Economy</td>
</tr>
<tr>
<td>NCHRD</td>
<td>National Center for Human Resources Development</td>
</tr>
<tr>
<td>PISA</td>
<td>Programme for International Student Assessment</td>
</tr>
<tr>
<td>QAD</td>
<td>Quality Assurance Directorate, Ministry of Education</td>
</tr>
<tr>
<td>SSME</td>
<td>Snapshot of School Management Effectiveness (SSME)</td>
</tr>
<tr>
<td>SPPD</td>
<td>Strategic Planning and Policy Directorate, Ministry of Education</td>
</tr>
<tr>
<td><em>Tawjihi</em></td>
<td>general secondary education examination</td>
</tr>
<tr>
<td>TIMSS</td>
<td>Trends in International Mathematics and Science Study</td>
</tr>
<tr>
<td>UNDP</td>
<td>United Nations Development Programme</td>
</tr>
<tr>
<td>UNRWA</td>
<td>United Nations Relief and Works Agency</td>
</tr>
<tr>
<td>USAID</td>
<td>United States Agency for International Development</td>
</tr>
</tbody>
</table>
About the Series

Building strong education systems that promote learning is fundamental to development and economic growth. Over the past few years, as developing countries have succeeded in building more classrooms and getting millions more children into school, the education community has begun to actively embrace the vision of measurable learning for all children in school. However, learning depends not only on resources invested in the school system, but also on the quality of the policies and institutions that enable their use and on how well the policies are implemented.

In 2011, the World Bank Group launched Education Sector Strategy 2020: Learning for All, which outlines an agenda for achieving “Learning for All” in the developing world over the next decade. To support implementation of the strategy, the World Bank commenced a multiyear program to support countries in systematically examining and strengthening the performance of their education systems. This evidence-based initiative, called SABER (Systems Approach for Better Education Results), is building a toolkit of diagnostics for examining education systems and their component policy domains against global standards, best practices, and in comparison with the policies and practices of countries around the world. By leveraging this global knowledge, SABER fills a gap in the availability of data and evidence on what matters most to improve the quality of education and achievement of better results.

SABER-Student Assessment, one of the systems examined within the SABER program, has developed tools to analyze and benchmark student assessment policies and systems around the world, with the goal of promoting stronger assessment systems that contribute to improved education quality and learning for all. To help explore the state of knowledge in the area, the SABER-Student Assessment team invited leading academics, assessment experts, and practitioners from developing and industrialized countries to come together to discuss assessment issues relevant to improving education quality and learning outcomes. The papers and case studies on student assessment in this series are the result of those conversations and the underlying research. Prior to publication, all of the papers benefited from a rigorous review process, which included comments from World Bank staff, academics, development practitioners, and country assessment experts.

All SABER-Student Assessment papers in this series were made possible by support from the Russia Education Aid for Development Trust Fund (READ Trust Fund). READ Trust Fund is a collaboration between the Russian Federation and the World Bank that supports the improvement of student learning outcomes in low-income countries through the development of robust student assessment systems.

The SABER working paper series was produced under the general guidance of Claudia Costin, Senior Director, Amit Dar, Director, and Harry Anthony
Osamha Obeidat and Zaina Dawani

Patrinos, Practice Manager in the Education Global Practice of the World Bank. The Student Assessment papers in the series were produced under the technical leadership of Marguerite Clarke, Senior Education Specialist and SABER-Student Assessment Team Coordinator in the Education Global Practice of the World Bank. Papers in this series represent the independent views of the authors.
About the Authors

Osamha Obeidat is the Inspection Program Manager at Abu Dhabi Education Council. Before joining the Council and when this case study was written, Osama was the monitoring and evaluation manager for the Queen Rania Foundation in Jordan. Between 2003 and 2009, he worked as an Associate Professor of Education at the Hashemite University in Jordan. Over the past 10 years, Osama has consulted for several local and international organizations to design monitoring and evaluation frameworks for organizations and programs and as an evaluator for education programs. His most recent consultation was with the Center for British Teachers (CfBT), where he led the development of curriculum standards for the Ministry of Education in Oman. Osama has received several scholarships, fellowships, and grants from the government of Jordan, universities in the United States, and from the EU through the TEMPUS mobility program. Osama has a Ph.D. in international and development education from the University of Pittsburgh and a certificate in evaluation practices from the George Washington University.

Zaina Dawani is an education consultant in the World Bank’s Middle East and North Africa region. She has over 10 years of regional experience in research, project management, and consulting, and was involved in education reforms in Jordan, Iraq, and Yemen. In Jordan, her work focused on research and assessment. She has led a number of studies, including those dealing with workforce development, benchmarking assessment, impact of teacher policies on the quality of teaching and learning, community colleges, the impact of the Tawjihi on education pathways, and focus group research on student aid. She has worked closely with educators in various capacities. Prior to joining the World Bank she worked for the United Nations Children's Fund (UNICEF) and United Nations Educational, Scientific, and Cultural Organization (UNESCO) in Jordan, managing education projects. Zaina is a Jordanian national and holds a Master’s degree in education and international development from the Institute of Education, University of London.
Acknowledgments

Many people provided inputs for this case study. We wish to thank Eng. Firyal Aqel, Development Coordination Unit, Ministry of Education, Hashemite Kingdom of Jordan; Ms. Miesser Al-Hindi, Directorate of Tests and Examinations, Ministry of Education, Hashemite Kingdom of Jordan; Dr. Khattab Abulibdeh, National Center for Human Resources Development; and Dr. Imad Ababneh, National Center for Human Resources Development. Thanks also are extended to the peer reviewers, Husein Abdul-Hamid and Michael Trucano, and to the members of the SABER-Student Assessment and READ teams, specifically Marguerite Clarke, Chelsea Coffin, Julia Liberman, Diana Manevskaya, and Maria Jose Ramirez. Special thanks to the Russian government for its support of this work under the READ Trust Fund program.
Abstract

Student assessment has been increasingly recognized as playing a key role in monitoring and improving student learning and achievement levels. This case study discusses the important steps that Jordan has taken to strengthen its student assessment system and ensure that the information gathered from classroom assessments, examinations, and large-scale assessment activities supports student learning. Countries seeking to strengthen the utilization and dissemination of student assessment data can extract the following lessons from Jordan’s experience: they should (i) set clear objectives for their assessment system, based on a well-defined vision of what is being assessed and why; (ii) entrust an external entity with managing system-level assessments in order to ensure impartial evaluation of the strengths and weaknesses of the education system; and (iii) introduce change gradually, coupled with adequate training on new assessment strategies and methods.
Executive Summary

This case study discusses the different uses of student assessments in the Jordanian education system. The purpose of the paper is to describe how Jordan makes effective use of its assessment information and to draw lessons for other countries that seek to improve their use of such data.

Student assessment has been increasingly recognized as playing a key role in monitoring and improving student learning and achievement levels. Jordan has taken important steps to strengthen its student assessment system. It has gradually built a hybrid system that encompasses the four main types of assessments found in the majority of education systems worldwide: examinations, classroom assessment, national large-scale assessments, and international large-scale assessments. Jordan’s student assessment system has evolved from consisting almost entirely of two elements: an examination that aimed to certify and select students (Tawjihi) and traditional classroom assessment activities. Today, its system encompasses updated and forward-thinking classroom assessment practices, as well as national and international large-scale assessments that influence the direction of education reforms in the country.

Currently, the use of assessment information varies depending on the type of assessment. The Tawjihi continues to be used to certify students for secondary education and admit them to postsecondary institutions; to some extent, it is also used to hold schools accountable for student performance. Classroom assessments are used to assist in the learning process (i.e., diagnose student learning problems), inform students and parents, and track students at the classroom level. Large-scale national and international assessments are used to inform policy making and education reforms, as well as to measure the impact of policy interventions.

There are three main drivers of student assessment in Jordan: the country’s National Agenda, its education reform program, and international organizations and donors. The Ministry of Education (MoE) began reforming the education sector in the early 1990s. However, the most comprehensive reform program was launched in 2003: the Education Reform for Knowledge Economy (ERfKE). This program has played a key role in shaping the student assessment system. ERfKE targeted all education levels, from basic and secondary education to vocational training and nonformal education, as well as governance and strategic planning in educational institutions. Improving the ways that learning outcomes are measured was one of the main interventions of the teaching and learning component of ERfKE. Through the monitoring and evaluation (M&E) framework developed by ERfKE, the country moved from implementing ad-hoc assessment activities to establishing a sustained, systematic approach to student assessment.
Other countries seeking to strengthen the utilization and dissemination of student assessment data can extract the following lessons from Jordan’s experience. These countries should:

- Set clear objectives for their assessment system, based on a well-defined vision of what is being assessed and why;
- Entrust an external entity to manage system-level assessments so as to ensure impartial evaluation of the strengths and weaknesses of the education system; and
- Introduce change gradually, coupled with adequate training on new assessment strategies and methods.
Disseminating and Using Student Assessment Information in Jordan

Osamha Obeidat and Zaina Dawani

Introduction

While many developing countries have managed to set up assessment systems, the effective use and dissemination of assessment data remains challenging for most. Jordan’s experience in developing a hybrid student assessment system provides useful insights and lessons for countries that are seeking to develop such systems and strengthen the use of assessment data in their education systems.

In 1962, Jordan instituted the general secondary examination (the Tawjihi) as the main national assessment. The exam had a dual purpose of certifying secondary school students and admitting students to postsecondary institutions. The Tawjihi became highly institutionalized and mandated by law; until recently, it was the education system’s only national assessment. With the introduction of education reforms in the early 1990s, Jordan began to prioritize assessments of learning outcomes that could support the learning process instead of simply certify or select students. The need for new types of student assessments increased after the launch of the Education Reform for Knowledge Economy (ERFKE) program in 2003. Currently, in addition to the Tawjihi, Jordan administers national large-scale and classroom assessments, and participates in the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA)—two well-known international assessments.

This case study examines how information from the student assessment system in Jordan is used and disseminated to improve education quality and learning. It is based on a review of reports and policy documents, together with meetings and interviews with key informants from the main assessment institutions of Jordan.

Country context

The Hashemite Kingdom of Jordan is a small, young Arab country that is strategically located in the Middle East. It is an upper-middle-income country with medium human development (according to the United Nations Human
Development Index), and is considered an emerging market. Jordan’s limited natural resources and weak industrial base make human capital a prime driver for economic development in the country. Between 2000 and 2008, Jordan enjoyed rapid economic growth that was higher than the regional average for the Middle East and North Africa. The global financial slowdown of 2008 and subsequent political unrest in the region, known as the “Arab Spring,” limited the country’s short-term growth prospects and increased unemployment. Jordan has a young demographic profile and one of the highest population growth rates in the world: 70 percent of its citizens are below the age of 30.

Over the past 20 years, the country has established a good track record of structural reforms across a number of areas. Reforms have progressed particularly well in education and training, as Jordan has placed a high emphasis on and made significant investments in improving the quality of its education system. These investments have had a noticeable impact on student learning since the early 1990s as evidenced by the trends in the country’s performance on TIMSS during this time period. In fact, between 1999 and 2007, no other country improved as much in its TIMSS science scores as did Jordan (World Bank 2009). Yet after 2009, Jordanian students’ performance on international assessments seriously declined, pushing the country to put a higher emphasis on its student assessment system.

Overview
This paper presents an overview of the student assessment system in Jordan. It describes the policy context in which the system has developed and how the system and its uses have evolved over the years with the introduction of education reforms. Four main types of assessments are discussed in the subsequent sections: (i) classroom assessment, (ii) examinations, (iii) national large-scale assessments, and (iv) international large-scale assessments. For each type of assessment, the paper outlines its uses and dissemination at different levels of the education system and government. It also discusses the drivers for change and offers key lessons and recommendations for countries seeking to improve their use of assessment data.

---

1 The Human Development Index is a summary measure used by the United Nations Development Programme (UNDP 2011) to assess long-term progress in three basic dimensions of human development: a long and healthy life, access to knowledge, and a decent standard of living.
1. Overview of the Student Assessment System

In order to understand the student assessment system in Jordan, it is important to have a sense of the broader policy context in which it developed. The student assessment system has undergone significant reforms in the last 30 years. Traditionally, the concept of classroom assessment in Jordan was confined to pencil-and-paper tests. These tests were typically administered several times per semester and influenced a student’s final course grades. Prior to the 1990s, the Tawjihi was the only national assessment that existed in Jordan. First implemented in 1962, it was used to certify secondary school students and admit secondary graduates to postsecondary institutions. Tawjihi results were also used by authorities and the public to evaluate schools. For example, parents used Tawjihi results published in the media to select the schools that they wanted their children to attend.

In light of the growing body of research on student assessments, which shows that adequate use of appropriate assessment activities contributes to improved learning outcomes and policy decisions (Heubert and Hauser 1999), policy leaders in Jordan determined that the country needed to adopt a systematic, sustained approach to student assessment beyond the Tawjihi. This system would complement reforms designed to transform the education system. Starting in 1991, Jordan began to participate in international large-scale assessments, with the purpose of benchmarking the country’s performance against that of other countries. It first participated in the 1991 International Assessment of Educational Progress II (IAEP-II). Since then, Jordan has participated in the TIMSS assessments of 1999, 2003, 2007, and 2011, and in the PISA assessments of 2006, 2009, and 2012.

After gaining experience in international large-scale assessments, Jordan developed and administered national large-scale assessments, including the National Test for the Control of Education Quality. The purpose of this new test was to monitor education quality in grades 4, 8, and 10, as well as to provide input for policy decisions. The first National Test for the Control of Education Quality was administered in 2000; the test was then radically reformed in 2004. However, this assessment has not been standardized, as assessment tools changed over time and its administration in each of the three grades was irregular. Moreover, the National Test only provided information on student achievement in certain subject areas and grade levels for a particular school year. Thus, in addition to periodically conducting the National Test for the Control of Education Quality, it was necessary to introduce other tools to measure trends in learning outcomes over time.
The introduction of the ERfKE I program in 2003 transformed the student assessment system in Jordan;² the program introduced curriculum reform as one part of transforming the country’s education programs and practices and sought to achieve learning outcomes relevant to the knowledge economy. The new curriculum defined three levels of learning outcomes: curriculum-specific, subject-specific, and class-specific. ERfKE also clearly outlined the knowledge economy skills that Jordanian students should possess upon completion of basic and secondary education. These skills included conventional academic skills; soft skills, such as communication skills; and personal management skills, such as responsibility, adaptability, and teamwork. It became the responsibility of the education system to build and nurture these higher-order skills. After identifying the outcomes and skills that students should learn, it was necessary to align the assessment system by identifying relevant ways to measure learning outcomes and assess the capacity of students to apply their knowledge and skills to real-life situations.

ERfKE I played a critical role in developing new assessment strategies through a monitoring and evaluation framework that included the measurement of learning outcomes. ERfKE’s framework required a sustained student assessment system that allowed for regular monitoring of learning outcomes at different levels of schooling (table 1). It included both internal monitoring (primarily related to the ERfKE project and its components) by departments of the Ministry of Education (MoE) and external monitoring and evaluation by the National Center for Human Resources Development (NCHRD) to ensure a neutral and impartial perspective (see figure 1 for a summary of the assessment institutions and entities in Jordan). NCHRD is responsible for carrying out all external evaluation activities for ERfKE, including external studies and assessments beyond the mandate of the Directorate of Examinations and Tests of the MoE and the M&E Division (MED) of its Quality Assurance Directorate (QAD).


The reform had two phases. Phase I focused on four main components: (i) re-orienting education policy objectives and strategies through governance and administrative reform, (ii) transforming education programs and practices to achieve learning outcomes relevant to the knowledge economy, (iii) supporting the provision of quality physical learning environments, and (iv) promoting learning readiness through early childhood education. Phase II focused on: (i) establishment of a national school-based development system, (ii) monitoring and evaluation, together with organizational development, (iii) development of teaching and learning, (iv) development of special-focus programs, and (v) improvement of physical learning environments.
<table>
<thead>
<tr>
<th>Indicator</th>
<th>Type of M&amp;E/evaluation</th>
<th>Organization responsible for data collection/study</th>
<th>Frequency of data collection</th>
<th>Possible approach to collecting data/methodology</th>
<th>Policy issues to be addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Increase in scores on national assessments aligned with knowledge economy skills (grades 5, 9, and 11 in math, science, and Arabic)</td>
<td>External M&amp;E (summary evaluation)</td>
<td>NCHRD obtains data from National Assessment for Knowledge Economy (NAfKE)</td>
<td>Report of NCHRD submitted to Directorate of Research and Planning (DRP), Strategic Planning and Educational Policies Analysis Division of the Strategic Planning and Policy Directorate (SPPD) of DRP identifies policy issues.</td>
<td>Years 2, 4, and 6 of ERfKE II Project (NAfKE test years)</td>
<td>Data collected from NAfKE tests further analyzed by categorizing them into rural and urban schools. Gender differences in scores are also analyzed. 1. If there are significant differences in scores between rural and urban schools, the reasons and corrective policies are elaborated, such as the allocation of more resources to lagging schools. 2. If there are significant differences in scores between male and female students, the reasons and appropriate corrective policies are elaborated.</td>
</tr>
<tr>
<td>Indicator</td>
<td>Type of M&amp;E/evaluation</td>
<td>Organization responsible for data collection/study</td>
<td>Organization responsible for dissemination of results</td>
<td>Frequency of data collection</td>
<td>Possible approach to collecting data/methodology</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>2. Enrollment rates: maintain current high net enrollment rates in basic cycle, and increase net enrollment rates in secondary cycle</td>
<td>Internal M&amp;E (combination of formative and summary evaluation)</td>
<td>Education Research and Studies Division (ERSD) of Education Research and Studies Directorate (ERSDir) of DRP will directly access the data from Jordan’s EMIS (once it is linked to the server). Until then, data is accessed through the EMIS Division (EMISD) of SPPD under DRP (based on the newly proposed organizational structure).</td>
<td>MED will organize workshop to disseminate results. In addition, quarterly meetings of M&amp;E Steering Committee will be held to provide a forum to study the external and internal evaluation reports of NCHRD and ERSD (under ERSDir) of DRP, respectively.</td>
<td>Annual</td>
<td>Secondary data assembled through EMIS. Both gross and net enrollment rates are needed to study coverage and efficiency of the school system. Further analysis will categorize the data into rural and urban schools and by gender.</td>
</tr>
<tr>
<td>Indicator</td>
<td>Type of M&amp;E/evaluation</td>
<td>Organization responsible for data collection/study</td>
<td>Organization responsible for dissemination of results</td>
<td>Frequency of data collection</td>
<td>Possible approach to collecting data/methodology</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>3. Programme for International Student Assessment (PISA)</td>
<td>External M&amp;E (summary evaluation)</td>
<td>NCHRD</td>
<td>NCHRD and DRP will organize workshops for all directorates of MoE, including field directorates. School principals and members of the Student Teacher Associations will also be invited.</td>
<td>PISA years (2012 and 2015), using 2009 as the baseline</td>
<td>Test results need vigorous analysis. Apart from comparing data to that of other countries, the analysis by NCHRD should reveal gender differences in scores of Jordanian students and also differences in scores between rural and urban schools.</td>
</tr>
</tbody>
</table>
### 4. Trends in International Mathematics and Science Study (TIMSS)

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Type of M&amp;E/evaluation</th>
<th>Organization responsible for data collection/study</th>
<th>Frequency of data collection</th>
<th>Possible approach to collecting data/methodology</th>
<th>Policy issues to be addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>External M&amp;E (summary evaluation)</td>
<td>NCHRD</td>
<td>NCHRD and DRP will organize workshops for all directorates of MoE, including field directorates. School principals and members of Student Teacher Associations will also be invited.</td>
<td>TIMSS years (2011 and 2015), using 2007 as the baseline</td>
<td>Test results need vigorous analysis. Apart from comparing data to that of other countries, the analysis by NCHRD should reveal gender differences in the scores of Jordanian students, as well as differences in scores between rural and urban schools.</td>
<td>1. The analysis recommended in the preceding column will give invaluable insights on gender bias (if it exists). 2. If there are significant differences in scores between rural and urban schools, policies such as improved resource allocation to hire better teachers should be implemented to help lagging schools.</td>
</tr>
</tbody>
</table>

**Note:** DRP — Directorate of Research and Planning; EMIS — education management information system; ERSD — Education Research and Studies Division; ERSDir — Education Research and Studies Directorate; M&E — monitoring and evaluation; MED — M&E Division, Quality Assurance Directorate, MoE; MoE — Ministry of Education; NAfKE — National Assessment for Knowledge Economy; NCHRD — National Center for Human Resources Development; PISA — Programme for International Student Assessment; QAD — Quality Assurance Directorate, MoE; SPPD — Strategic Planning and Policy Directorate; Tawjihi — general secondary education examination; TIMSS — Trends in International Mathematics and Science Study.
**Figure 1. Assessment Institutions and Entities in Jordan**

**MINISTRY of EDUCATION**

- Managing Directorate of Examinations and Tests
  - Directorate of Tests
  - Directorate of Examinations
  - M&E Division
  - EMIS Division

- Managing Directorate of Planning and Research

**Central Level**

- Education Directorate
  - Governorate Division of Directorate of Examinations & Tests
  - Governorate Division of Directorate of Planning & Research

**Governorate Level**

**KEY: Ministry of Education**

**Directorate of Examinations and Tests of MoE**: The managing Directorate of Examinations and Tests at MoE has two directorates: the Directorate of Tests and the Directorate of Examinations. The Directorate of Tests is responsible for administering the National Tests for Control of Education Quality, as well as for designing and guiding classroom assessments. The Directorate of Examinations is responsible for administering the Tawjihi examination.

**Directorate of Examinations and Tests at the governorate level**: In each governorate (city) in Jordan there is an Education Directorate that houses the main divisions of MoE, including the Directorate of Examinations and Tests. The governorate division of the Directorate of Examinations and Tests is responsible for implementing and monitoring all activities associated with tests and examinations in the governorate. This division has two types of staff: supervisors in the Education Directorate and field or district supervisors mapped to 42 districts.

**M&E Division (MED) and EMIS Division of Directorate of Planning and Education Research of MoE**: MED is responsible for internal M&E of ERfKE, and the EMIS Division is responsible for monitoring both Jordan’s education management information system and all education indicators. Both divisions are under the same directorate and coordinate closely.

**National Center for Human Resources Development (NCHRD)**: The NCHRD is a semi-governmental body that was created with the mandate to conduct assessments and lead research studies that inform policy making on education. It is entrusted, among other things, with conducting the external evaluations of ERfKE, the National Assessment for Knowledge Economy, TIMSS, and PISA.
In 2009, the second phase of ERfKE was launched. Again, student assessment was one of the main interventions under its Development of Teaching and Learning component (World Bank 2009). In particular, ERfKE II focused on refining and developing tools and training for assessment activities at both the classroom and national level. As a result, the monitoring and evaluation framework continued to be revised under ERfKE II, with a new organizational structure proposed for the monitoring and evaluation division of the Ministry of Education. The program also elaborated the roles and responsibilities of internal and external parties in monitoring learning outcomes, including NCHRD. The monitoring and evaluation framework helped the MoE regularly measure whether its programs were achieving their goals and identify any bottlenecks during implementation. In a recent study by World Education, Inc. and NCHRD on the relevance and utilization of NCHRD’s external evaluations of ERfKE II, most participants from the Ministry of Education staff expressed positive views on the usefulness of these studies in supporting the development and revision of programs implemented by the MoE (WEI and NCHRD 2013).

In addition to systemizing participation in international assessments, such as TIMSS and PISA, and improving classroom assessment strategies, ERfKE I also introduced a new national assessment, the National Assessment for Knowledge Economy (NAfKE). The latter aimed to identify trends in education outcomes and measure the impact of ERfKE interventions in order to inform education policies and reforms. NAfKE measures students’ mastery of knowledge economy skills in math, science, and reading in grades 5, 9, and 11.

MoE continues to build on the progress achieved by both improving existing assessments and monitoring activities and further developing the student assessment system. In 2012, MoE in partnership with the United States Agency for International Development (USAID) introduced a new national assessment for the early grades, “The National Survey on Student Performance in Reading and Mathematics, Pedagogic Practice and School Management”. The National Survey consists of three components: (i) Early Grade Reading Assessment (EGRA); (ii) Early Grade Mathematics Assessment (EGMA); and (iii) Snapshot of School Management Effectiveness / (SSME). The Survey will be conducted again in 2014. Figure 2 presents the evolution of the student assessment system in Jordan.
2. Classroom Assessment

The concept of classroom assessment in Jordan was previously confined to pencil-and-paper tests that were generally administered a few times per semester and influenced students’ final course grades. However, a new understanding of classroom assessment has come into being in recent years. Assessments are now seen as more “formative” in nature, a change that conforms to new pedagogical methods that call for the use of alternative assessments associated with deeper student learning (Rust 2002). Today, classroom assessment in Jordan encompasses multiple components, such as scoring of students’ work, informal observations of students, self-reflection, quizzes, and end-of-year testing.
Box 1. Dissemination and Use of Classroom Assessment Information

Classroom assessment activities that promote and inform student learning include diagnosing student learning issues, providing feedback to students on their learning, informing parents about their child’s learning, and planning the next steps in instruction. Classroom assessment activities are also used to determine whether students have passed, failed, or should repeat a school year, as well as to determine the tracking of students (i.e., which stream students should pursue in secondary education—scientific, literary, information technology, etc.—as each stream has a minimum entrance grade).

Teachers typically carry out classroom assessment activities to inform their own teaching and their students’ learning, as well as to meet external (system-level) requirements or information needs. Classroom assessment activities focus on knowledge and skills in core subjects (such as reading and writing) and non-core (such as civics and home economics) curriculum areas, as well as on non-cognitive skills (such as team work and self-discipline).

Schools and teachers are required to communicate classroom assessment information to students and parents. A student’s classroom assessment results are typically recorded in the teacher’s record book as well as in the student’s own copy book. Classroom assessment results are also recorded in the classroom or the school database, as well as in a system-wide online database called EduWave. EduWave is an online platform that was established to support student learning. It allows students to track their progress, improve their school performance, and enjoy their learning experience. The system enables teachers to use interactive examples, models, and scientific experiments to enrich conventional textbook material. Schools also submit quarterly classroom assessment results to the Directorate of Education, which then submits them to the MoE.

System-level mechanisms set classroom assessment standards and criteria; these criteria monitor the quality of classroom assessment activities. Specifically, classroom assessment is a required component of a teacher’s performance evaluation, and of school monitoring or teacher supervision. Supervision is conducted by the Directorate of Examination and Tests, and field visits to schools are carried out to follow up on implementation of standards and criteria. At the end of each school term, the Directorate of Examinations and Tests prepares a report on identified issues, based on a sample of schools visited, along with recommendations.

The Ministry of Education mandates the design of classroom assessment activities, which have undergone several changes and developments over the years. The curriculum reform under ERfKE pushed to align school system components, including new assessments. In 2004, guided by the new curriculum and the ERfKE assessment framework, a comprehensive “Assessment Strategy 2004” (MoE 2004) was developed for classrooms that introduced new assessment tools and concepts. The strategy introduced “authentic assessment,” which uses five different methods to assess students at the classroom level (box 2) and aimed to transform the traditional system from “assessment of learning” to “assessment for learning” (MoE 2004).

Teachers are expected to use the methods shown in box 2 to conduct classroom assessment activities. Field supervisors from the Directorate of Examinations and Tests are responsible for training teachers on how to use and implement these classroom assessment methods during in-service teacher training. However, when the 2004 Assessment Strategy was first introduced, in-service teacher training was neither structured nor organized by the Ministry of Education. Instead, each field supervisor delivered training to teachers based on each teacher’s individual work plan.
Box 2. Methods of “Authentic Assessment”

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Performance-based assessment: students demonstrate how to use skills in real-life situations through simulations or presentations.</td>
</tr>
<tr>
<td>2.</td>
<td>Pencil-and-paper assessments: students take tests, quizzes, and exams.</td>
</tr>
<tr>
<td>3.</td>
<td>Oral assessments: students are tested via interviews or question-and-answer sessions in order to assess their reasoning and problem-solving abilities.</td>
</tr>
<tr>
<td>4.</td>
<td>Classroom observations: teachers observe students’ behavior and progress.</td>
</tr>
<tr>
<td>5.</td>
<td>Self-reflection: students diagnose their own strengths and weaknesses in order to identify their needs.</td>
</tr>
</tbody>
</table>

*Source: MoE (2004).*

Recently, the Directorate of Examinations and Tests of the Ministry of Education conducted a field study as part of its internal monitoring to assess how teachers use the different assessment methodologies in the classroom. The study revealed that teachers lacked the knowledge needed to use and implement these methodologies, and that the field supervisors who trained them possessed only 40 percent of the material and content knowledge needed to do so. In addition, the study showed that the monitoring and evaluation of the implementation of these methodologies by the Directorate of Examinations and Tests was weak.

Based on the findings of this study, that directorate formulated a new training plan for field supervisors and teachers. Its supervisors now train focal points from each of the 42 school districts; a supervisory team is also formed in each school district that consists of the head of supervisors, two scientific subject supervisors, and two humanities subject supervisors. The team trains and supervises the teachers in the district continuously, particularly the assistant principal, coordinators of scientific subjects, and coordinators of humanities subjects.

Finally, the 2004 Assessment Strategy served as the basis for other classroom assessment guidelines. To implement the strategy and improve classroom assessment activities, the following guidelines were developed: (i) brochures on classroom assessment (for grades 1–12), distributed to schools as a reference; (ii) an achievement test guide, which provides a general overview of how to design and conduct achievement tests; and (iii) performance standards for four core subjects (Arabic, English, science, and math) at all levels of schooling. These guidelines have helped improve classroom assessment practices and informed teachers and schools about their use. For example, teachers now develop their own individual assessment tools to measure the performance standards and indicators outlined in the guidelines. In addition, the achievement test guidelines helped schools design and implement their own achievement tests.

---

3 The source of this information is the director of tests at MoE’s Directorate of Examinations and Tests.
Along with formative classroom assessment activities, students are required to take four classroom-level tests during the school year, each of which contributes a percentage to the student’s final grade in a subject. The first, second, and third test each contribute 20 percent to the student’s final grade. The fourth test contributes 40 percent to the final grade. A pencil-and-paper format is used for the first, second, and fourth tests. Authentic assessment is employed as part of the first and entirely as the third test. However, teachers tend to confuse authentic assessment with student participation and activity in the classroom and grade students’ performance on authentic assessment as they understand it.

**Uses of classroom assessment data**

There are three main uses of classroom assessment information in Jordan:

1. **Diagnosing student learning issues (formative).** Teachers use the results of classroom assessment activities to inform their teaching process by judging student performance, identifying learning issues, and facilitating active student learning.

2. **Informing students and parents.** Classroom assessment administered results are shared with students and parents. The results of tests administered in the classroom, together with continuous feedback, are provided to students, giving them clear indications of their respective strengths and weaknesses. In addition, a “parent card” is provided to parents at the end of each school quarter. This card contains information on a student’s marks in different subjects, together with remarks from class teachers and head teachers on his/her general performance.

3. **Tracking students.** One of the critical uses of classroom assessments is to track students into specific education streams at the secondary level (e.g., scientific, literary, information technology, etc.). For each secondary stream, the entrance requirement is a minimum grade in relevant subjects in grade 9. Classroom assessment results also determine whether a student can move on to the next grade or must repeat a grade.

**3. General Secondary Examination (Tawjihi)**

The Tawjihi is a high-stakes examination that assesses students’ knowledge on the five mandatory subjects common to all streams of secondary education (Arabic, English, Islamic education, social studies, and computer science), as well as on optional subjects. The examination is administered twice a year (once a

---

4 The phonetic pronunciation of the Tawjihi is: Taw-je-hie. Secondary education in Jordan has five main academic streams (literature, science, information technology, Islamic studies, and health) and four vocational streams (industry, agriculture, home economics, and hotel management and tourism).
semester) to grade 12 students. The overall score is determined based on a student’s performance in each subject, thus his/her performance on each subject contributes a percentage to the overall score. The Tawjihi is fully standardized and centrally managed by the MoE.

### Box 3. Dissemination and Use of Tawjihi Data and Results

The Tawjihi has been administered in Jordan since 1962 to certify student completion of the school cycle and to determine selection to higher education institutions. Therefore, the Tawjihi remains the most important form of assessment for students and parents, as it determines students’ educational pathways. Views about the high-stakes uses of Tawjihi vary, with many opposing the nature of the examination. At the same time, there is no systematic evidence of improper use of examination results by any stakeholder group, and Tawjihi results are perceived as credible by all of them. Many measures are in place to guarantee the examination’s confidentiality, precise execution, and accurate scoring. Student results are not made publically available. Instead, students can access their results online using their unique seat number. The Tawjihi is deeply rooted in the education system in Jordan and is a major determinant for completion of secondary education and entrance to university. Focus groups and surveys of key stakeholders are conducted on a regular basis.

In addition to being a high-stakes exam, the Tawjihi is the assessment program with the greatest impact on students’ postsecondary education options. Nearly 139,000 students sit for the exam annually, but only 68,000 pass. Only those who pass the examination are eligible to apply to postsecondary educational institutions. Students who pass the Tawjihi can apply to either public or private postsecondary educational institutions. Private institutions have less rigid selection criteria, primarily a minimum grade on the Tawjihi, while admission to governmental institutions is based on more selective admission criteria and controlled by the Unified Admission Coordination Unit under the Higher Education Council. If students do not pass the Tawjihi on their first attempt, they are allowed to sit for the examination or the subject that they failed four additional times in the three years following their initial examination.

**Certifying and selecting students**

As noted earlier, the main purpose of the Tawjihi is to provide students with a secondary-school certification, based on examination results in the mandatory and optional subjects in their specific educational stream. The examination is also used to select students for admittance to postsecondary institutions. Students can apply to over 20 universities; students with higher scores have a better chance of admission into the universities or programs of their choice. Each year, the Unified Admission Coordination Unit sets a minimum score for admission to university programs based on the results of the Tawjihi, the caps set by each individual university, and the recommendations of the Higher Education Council. If a student’s score does not make him/her eligible for admittance to

---

5 Students must have a minimum GPA of 65 percent on the Tawjihi to be eligible to apply to public universities and a minimum of 55 percent to apply to private universities.
his/her preferred universities or programs, he/she is automatically admitted to a less selective university or program. Students who do not qualify for university programs are admitted to community colleges.

The Directorate of Examinations and Tests of the Ministry of Education issues the Tawjihi results and sends them to both the directorates of education in the governorates and individual schools. Students can access their Tawjihi results either through the “EduWave” online system by using their unique seat number, or through their school. Access to Tawjihi results is free of charge and provided in a confidential manner.

The secondary (although not formally intended) use of Tawjihi examination scores is school accountability. While there is no formal school accountability system in Jordan, parents and students attach high importance to Tawjihi results. Through these results, schools are held accountable to parents, students, and the education directorate in the relevant governorate. Schools and teachers whose students receive high Tawjihi scores are more likely to make their high scores known (without disclosing the identity of specific students) in order to build a good reputation in the community and therefore attract more students (students select the schools they attend), as well as to achieve recognition by authorities at all levels. Thus, teachers tend to teach to the Tawjihi very early in the students’ learning cycle, focusing solely on topics or academic skills that can help students achieve high scores on the exam.

Reforming the Tawjihi has always been on the education reform agenda in Jordan. Policymakers’ views of the Tawjihi have always conflicted; some oppose it, while others support it. Issues that have been raised include the examination’s high-stakes nature and its focus on recalling information. However, only procedural reforms have taken place to date, including improvements in the administration of the exam and the introduction of quality control mechanisms.

One of the subcomponents of ERfKE II is to reform the Tawjihi to ensure that it is aligned with the new curriculum and assessment framework. In 2013, a study was undertaken by the Ministry of Education to assess to what extent the exam measures knowledge economy skills. The study reviewed exams from 2010, 2011, and 2012 for the same subjects that are part of NAfKE, with negative results across all subjects (MoE 2013b). For example, the average weights given to knowledge economy skills in math across the three years were: 25 percent for communication skills, 68 percent for information management, and 7 percent for problem solving and critical thinking. The results of this study have not, however, led to concrete plans to adjust the misalignment between the new curriculum and the Tawjihi.
4. Large-Scale National Assessments

There are two national large-scale assessments in Jordan: the National Test for the Control of Education Quality and the NAfKE. The former was the first national large-scale assessment in the country, launched in 2000. This assessment originally measured the skills of students in grades 3 and 9 in six subjects (English, Arabic, math, science, Islamic education, and social studies) on a census basis. However, this test was reformed in 2004. Currently, it assesses students’ skills in four subjects (Arabic, English, math, and science). Every year, students from one grade out of grades 4, 8, and 10 are tested in the four subjects.

These changes and irregularities left gaps in the assessment data that made it difficult to ascertain learning trends. To fill these gaps, a new national assessment, NAfKE, was introduced over the course of both phases of ERfKE. NAfKE measures students’ mastery of knowledge economy skills in math, science, and reading based on a representative sample of students in grades 5, 9, and 11. It is one of the key indicators for measuring progress towards the development objectives of ERfKE. It is not yet clear whether NAfKE will be sustained after the reform.

In addition to these two main national assessments, Jordan, in partnership with USAID, recently introduced “The National Survey on Student Performance in Reading and Mathematics, Pedagogic Practice, and School Management.” The survey aims to gain insight into student performance in reading and mathematics as foundational and predictive skills, as well as to better understand the characteristics of Jordanian schools associated with this performance. The survey has three assessment components:

(i) **Early Grade Reading Assessment (EGRA):** designed to orally assess the most basic foundation skills for literacy acquisition in early grades, including pre-reading skills, such as listening comprehension.

(ii) **Early Grade Mathematics Assessment (EGMA):** an individually administered oral assessment of foundational mathematics skills. Administered on a sample basis with external assessors, it aims to provide a system snapshot of areas of skill deficiency.

(iii) **Snapshot of School Management Effectiveness (SSME):** an instrument that yields a quick, but rigorous, multifaceted picture of school management practices in a country or region. The resulting data are designed to let school, district, provincial, and/or national administrators learn what is currently going on in their schools and classrooms and to assess ways in which to make these schools more effective.

The National Survey was conducted for the first time in spring 2012, covering 156 public primary schools across Jordan, with the participation of 3,120 students randomly selected from grades 2 and 3. It included interviews with
teachers and school principals from the selected classes, as well as classroom observations of the selected grade 2 teachers while they taught a reading and a mathematics lesson. The survey will be implemented again in May 2014. Given that this assessment is new, the dissemination and use of its data are not yet clear; therefore this paper limits itself to the two main assessments mentioned above: the National Test for the Control of Education Quality and NAfKE.

The National Test for the Control of Education Quality
The “National Test” is a three-year cycle study. The test has undergone several reforms, most recently in 2012, when its focus shifted from assessing students at a given grade level to assessing students at an education level (cycle). The new test design is intended to evaluate what students learn in a full education cycle, based on performance indicators developed for that specific cycle in each subject. Scores on the National Test are represented by four achievement levels (table 2).

<table>
<thead>
<tr>
<th>Level</th>
<th>Score</th>
<th>Meaning of level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic</td>
<td>≤ 34%</td>
<td>Student does not meet minimum expected learning outcomes and needs a remedial plan to reorient his/her learning path</td>
</tr>
<tr>
<td>Nearly proficient</td>
<td>35% to 59%</td>
<td>Student meets minimum expected learning outcomes</td>
</tr>
<tr>
<td>Proficient</td>
<td>60% to 84%</td>
<td>Student meets expected learning outcomes</td>
</tr>
<tr>
<td>Advanced</td>
<td>≥ 85%</td>
<td>Student exceeds preset learning standards</td>
</tr>
</tbody>
</table>

The National Test is designed and administered by the Directorate of Examinations and Tests of the Ministry of Education. A high-level committee comprised of the undersecretary and heads of the curriculum, training, planning, and tests directorates is established within the MoE to make decisions and choose the strategic direction of the test. The committee endorses any changes in assessment tools, reviews the results before they are published, and provides regular feedback to the Directorate of Examinations and Tests. In addition, for each subject assessed by the National Test, a technical committee is convened to design the test and decide on changes to any of its parameters. This technical committee is comprised of measurement and evaluation supervisors and representatives from the curriculum, textbook, and training directorates.

Uses of assessment results
The National Test has two intended purposes: to help teachers identify students’ strengths and weaknesses and identify areas for improvement; and to provide policy makers with information about the quality of education, thus enabling

---

6 The cycles are as follows: grades 1 through 4, grades 5 through 8, and grades 9 and 10.
them to make informed decisions about education reform. While the dissemination strategy for assessment results is aligned with these purposes, the actual use of National Test results is confined to helping teachers.

**Box 4. Dissemination and Use of Data and Results from the National Test for the Control of Education Quality**

The National Test for the Control of Education Quality, administered by the Ministry of Education, was introduced as the first large-scale national assessment in 2000, and was radically reformed in 2004. The Directorate of Examinations and Tests disseminates the results of the test annually. A detailed report is published that includes assessment data and graphs at the school level. The findings are first shared with the Board of Education for discussion and finalization, and then with three different levels of the education system: directorates at the central and governorate level and with schools and teachers. Schools are asked to develop action plans addressing the weaknesses identified by the report.

Initially, as the name indicates, the main purpose of the test was to inform policy making by providing a clear picture of the quality of education at the national level. However, the design and uses of test results were not entirely consistent with the test’s intended purposes. The National Test now primarily provides pedagogical support to teachers, although this support is limited. Specifically, schools and teachers use specific information provided by the results report, such as the mean score of each school and student compared to district and national mean scores in all subjects and skills, to determine school performance and student abilities. While such information helps teachers compare performance levels, no further support or analysis is provided to help schools and teachers address the issues identified by the test.

Dissemination of test results occurs in two stages. First, the findings are published by the Directorate of Examinations and provided to the Ministry of Education and the Board of Education (comprised of several ministers, including the Ministers of Education, Labor, and Higher Education, plus high-level officials from the Ministry of Education) for discussion and finalization of a report on the test (figure 3). The Directorate of Examinations and Tests shares the final report with the Directorate of Curriculum and Textbooks; the Directorate of Training, Supervision, and Qualifications; the Directorate of Private Education; and the Directorate of Planning and Educational Research (formerly the Directorate of Research and Education Development). The purpose of this exercise is for these directorates and their respective policy makers to act upon the information provided in the report. In practice, however, the results are used only by the Directorate of Examinations to inform test design.
Once the report is disseminated to the four aforementioned directorates, it is distributed to school districts. Each of the 42 school districts receives one hard copy of the complete report and one CD with the results presented by district, school, and student level.

The National Test results provide pedagogical support to teachers and schools through such specific information as the mean score of each school and student compared to district and national mean scores in all subjects and skills (figure 4). However, the reports only present information in the form of graphs; no further analysis is provided. Each school district then forms a committee comprised of the supervisors of each subject for the district, which discusses the findings and possible future actions, particularly for the weakest schools within the school district. The school district disseminates the results to all schools. With the support of the supervisors of the relevant district, each school organizes a meeting with teachers, students, and the principal to discuss the results. The school is then asked to design, with the support of the supervisor, an action plan to address the weaknesses identified in the discussion. This action plan is sent to the Ministry of Education within one month of receipt of the National Test results.
Figure 4. Example of Schools Report, 2012–2013

While these reports help teachers and schools identify weaknesses and benchmark their performance to that of other schools and districts, they have limited pedagogical use because schools and teachers are not assisted in further analyzing and understanding assessment results. Teachers are also not supported in using the results to improve their teaching; therefore many of the action, or remedial, plans are not well informed and their implementation is weak.

**The National Assessment for Knowledge Economy (NAfKE)**

As previously explained, NAfKE was constructed to predict students’ abilities in a knowledge economy by measuring their academic readiness to utilize knowledge and concepts to solve problems in real-life situations. First conducted in 2006, its findings were the baseline for ERfKE I. NAfKE was also conducted in 2008 and 2011 to measure progress in students’ achievement levels. The 2011 NAfKE results served as the baseline for ERfKE II.

---

**Box 5. Dissemination and Use of Data and Results from the NAfKE**

The NAfKE is conducted by NCHRD alongside the Education Reform for Knowledge Economy (ERfKE I and II) projects to measure progress towards their development objectives. In addition, NAfKE is used to fill the gaps in existing tools and assessments to measure trends in the education system. NAfKE results are disseminated within 12 months after the large-scale assessment is administered, via reports that target different stakeholders. For example, the technical report targets technical specialists, while policy briefs are developed for policy makers and ERfKE partners. Reports with results contain information on overall achievement levels and by subgroups, trends over time, as well as standard errors. Workshops are also organized to communicate NAfKE results to stakeholders.

Initially, the use of NAfKE results were not entirely consistent with the test's stated purposes (because neither the 2006 nor the 2008 study was seen as a baseline against which to measure progress); however, after recent results indicated a decline in student performance, the MoE led the development of a comprehensive action plan to holistically address the decline. The action plan addressed topics from curriculum revision and awareness raising on key curricular topics at the school level to teacher training and increasing time allocated to teachers for teaching certain subjects.
NAfKE is administered by NCHRD, which coordinates with the Directorate of Examinations and Tests of the Ministry of Education to form a committee of representatives from three authorities: the Ministry of Education, the United Nations Relief and Works Agency (UNRWA), and the schools administered by the Ministry of Defense. In addition, a coordinator is appointed to manage enumerators and field researchers in each school district. The NAfKE study sample is chosen from the school population as a stratified random cluster in two stages: (i) schools are chosen from the study population; and (ii) the class in the school is chosen.

**Policy interventions and their impact**

The primary use of NAfKE results is to measure the strengths and weaknesses of students in knowledge-economy skills before and after the implementation of ERfKE. The test data informs relevant decision making and education reforms. The ERfKE results framework comprises a number of studies and assessments, with NAfKE being a major input into the evaluation of its impact.

Following the administration of NAfKE, NCHRD conducts data analysis and prepares a detailed report that describes the technical aspects of the assessment and analyzes the results. NCHRD prepares both policy briefs (figure 5) and a detailed report (figure 6); the results are then shared with decision makers at different levels. The report is shared in a half-day meeting with the Secretary General of the Ministry of Education and the heads of planning, curriculum, training centers, the Directorate of Examinations and Tests, the Development Coordination Unit (DCU) of the Ministry of Education, and all ERfKE partners. The DCU ensures that issues identified in the report are incorporated into the ERfKE program and raised with partners during ERfKE supervision missions.
Figure 5. NAfKE Policy Brief, 2013

The decline ranged between (2-18) scores and all declines were statistically significant. The decline size average increased by grades' levels five, nine, then eleven. It also varied by domains as the highest was in literacy, then science, and the lowest performance was in Math regardless of the grades levels.

Approach and findings
Jordan has always been keen on conducting national studies and participating in international ones to get a hold of indicators that support the assessment of the educational system through identifying the strengths to enhance them and the weaknesses to develop appropriate plans to address them. The National Assessment for Knowledge Economy Skills (NAfKE) is one of the national studies conducted abroad before both phases of Education Reform for Knowledge Economy (ERfKE I & II). This study was first conducted in 2006, and its findings were the baseline for ERfKE I. The study was again conducted in 2008 and its findings showed marked progress and improvement in students’ compared to 2006. In addition, the study was conducted for the third time in 2011 and the findings showed that there is apparent decline in students’ performances compared to 2006 and 2008. An in-depth report on the decline factors / 2011 was conducted as a serious attempt to reveal the factors that led to this decline. It is also worth mentioning that the findings of the International Study on “Trends in Math and Science /TIMSS 2011” revealed that there is a statistically significant decline compared to the 2007 findings, and the decline size was (33) scores in science and (21) scores in Math.

To measure the decline size, the difference between students’ performance averages in 2011 and in 2008 was calculated by grade and by domain as the decline ranged between (2-8) at the country’s level. All declines were statistically significant at (α = 0.01), and the average decline size increased by grade level as the highest decline size was at grade 11 students. Moreover, the decline size varied by domain as the highest decline was in the literacy domain, and the lowest was in the math domain, while the decline size in the science domain ranked between the math domain and literacy domains. In 2008, Students’ performance was the lowest in the math domain and the highest in the literacy domain regardless of the grades. In 2011, students’ performance was the lowest in the math domain for all grades, and the highest was in the literacy domain for grades (9&11), while it was the highest in the science domain for grade (5). Regarding gender, school location and the supervising authority variables, the decline size differed. The findings indicated that the decline size was higher at males compared with females as well as students in rural areas compared to students in urban areas. For the supervising authority, the decline size was the highest at students in public schools and the lowest at students in UNRWA schools.

Furthermore, the findings highlighted that the decline size varies by percentiles, as the decline size increased by the student’s achievement ability. This might indicate that the education’s attention to the students with low achievement is higher than that to students with high achievement.

Figure 6. NAfKE Results, 2011: Achievement Levels by Subject and Grade

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Advanced</td>
<td>High</td>
<td>Intermediate</td>
<td>Low</td>
<td>Very Low</td>
<td>Below</td>
</tr>
<tr>
<td>elementary</td>
<td>%1</td>
<td>%1.2</td>
<td>%4.1</td>
<td>%17.7</td>
<td>%76.1</td>
<td></td>
</tr>
<tr>
<td>advanced</td>
<td>%0.3</td>
<td>%6.4</td>
<td>%8.8</td>
<td>%25.8</td>
<td>%48.7</td>
<td>%13.0</td>
</tr>
<tr>
<td>first</td>
<td>%1.2</td>
<td>%0.4</td>
<td>2.8</td>
<td>%11.2</td>
<td>%30.9</td>
<td>%53.4</td>
</tr>
<tr>
<td>second</td>
<td>%1.1</td>
<td>%11.4</td>
<td>%26.6</td>
<td>%30.5</td>
<td>%18.0</td>
<td>%12.4</td>
</tr>
<tr>
<td>third</td>
<td>%0.6</td>
<td>%6.1</td>
<td>%15.7</td>
<td>%28.1</td>
<td>%30.1</td>
<td>%19.4</td>
</tr>
<tr>
<td>fourth</td>
<td>%0.1</td>
<td>%3.0</td>
<td>%10.5</td>
<td>%18.0</td>
<td>%31.7</td>
<td>%36.7</td>
</tr>
<tr>
<td>fifth</td>
<td>%1.2</td>
<td>%9.0</td>
<td>%19.1</td>
<td>%25.8</td>
<td>%26.9</td>
<td>%18.0</td>
</tr>
<tr>
<td>sixth</td>
<td>%0.3</td>
<td>%3.4</td>
<td>%18.7</td>
<td>%33.5</td>
<td>%26.3</td>
<td>%13.7</td>
</tr>
<tr>
<td>seventh</td>
<td>%1.8</td>
<td>%10.5</td>
<td>%32.8</td>
<td>%34.9</td>
<td>%16.0</td>
<td>%6.8</td>
</tr>
<tr>
<td>eighth</td>
<td>%0.8</td>
<td>%2.2</td>
<td>%34.2</td>
<td>%34.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: NCHRD (2011b).

Uses of assessment results

The 2008 NAfKE results showed marked improvement in student learning achievement compared to the baseline of 2006. At the time, use of these results was confined to producing evidence of the positive impact of ERfKE. However, the 2011 NAfKE results indicated a decline in student performance compared to both 2006 and 2008. The Ministry of Education and NCHRD determined that further investigation and analysis were needed to determine why, and an additional detailed study was undertaken by NCHRD. All of the issues identified were fed back into the design of ERfKE II through the DCU, particularly with respect to curriculum revision, teacher training, and classroom assessment components (box 6).
Box 6. Factors Contributing to Decline in Student Achievement on NAfKE

Teacher-related factors

- Overloaded teaching burden
- Low social status of teachers
- Crowded classrooms
- Lack of effective training for education innovations introduced by Ministry of Education
- Poor student discipline
- Lack of parental engagement
- Lack of required facilities
- Low level of student interest in education
- Student behavioral problems
- Lack of student desire and motivation to learn

School-related factors

- Lack of maintenance for school buildings
- Lack of education resources
- Damage to school properties
- Teachers’ preoccupation with work other than teaching
- School environment
- Lack of teacher motivation
- Teachers’ absence
- Decline in achievement varied by gender and the school location (rural, urban), as well as by the supervising authority (MoE, Military Education, UNRWA, private educational institutions)

Possible Remedial Policies

- Gradually recruit female teachers in the early basic education cycle
- Conduct twinning projects between boys and girls schools in the upper basic and secondary cycle to exchange experience as well as disseminate success stories
- Focus on the early basic education cycle, especially in math and Arabic
- Focus on the organizational environment of schools and their values systems to reduce school discipline problems and build healthy relationships between each school and its local community.


5. International Assessments

Jordan has participated in international large-scale assessments for the purpose of benchmarking its performance against countries in the Middle East region and worldwide; specifically, against those in the Middle East region. With the introduction of education reforms in the early 1990s (World Bank 2009), it became clear to policy makers that measuring whether students were meeting international learning standards was the key to improving the quality of education in Jordan. Understanding what students outside of Jordan were learning, and whether Jordanian students were learning 21st-century skills that would enable them to join the global economy, could be answered by their participation in international assessments and comparing their scores with those of students from other countries. The first several rounds of Jordan’s
participation in international assessments were not systematically planned. The country did, however, recognize that participation in these assessments provided insight into how Jordanian students compared to students in other countries in key subjects over time, as well as provided contextual information that helped the country make informed decisions on how to move forward in improving its education system.

Jordan participated in its first international large-scale assessment, the Second International Assessment of Educational Progress (“IAEP-II”), in 1991 as one of 20 participating countries. The purpose of the study was to measure the knowledge of 13-year-old students in mathematics and science. A nationally representative sample of students in grade 8 was assessed in the two subjects. As of 1999, Jordan had participated in international assessments on a regular basis. Under the leadership of NCHRD, which is responsible for administering all such assessments in the country, the country has now participated in eight such assessments, including TIMSS grade 8 (1999, 2003, 2007, and 2011) and PISA (2006, 2009, and 2012). Students from 210 schools in Jordan, including urban, rural, private, UNRWA, and Ministry of Defense schools, participated in each study.

Box 7. Dissemination and Use of International Assessment Data and Results

In the last ten years, Jordan has participated in the Second International Assessment of Educational Progress (IAEP-II) (1991), TIMSS (1999, 2003, 2007, 2011), and PISA (2006, 2009, 2012). The results of these assessments are presented in detailed national and international reports targeted to different stakeholders, such as the MoE and donors. A policy brief targeting policy makers and other key stakeholders is also prepared. Because there was an apparent decline in the most recent TIMSS 2011 results compared to previous years (NCHRD, 2011a), another report was produced to analyze results as well as decline factors. The reports are published online, distributed to key stakeholders, and communicated through press releases and Power Point presentations. The database with international large-scale assessment results, the national and international reports, and the policy brief are made available on a CD that is disseminated to universities and other interested institutions such as various MoE departments and other relevant ministries. Media coverage of international large-scale assessment results has been inconsistent from year to year. For example, while the TIMSS 2003 results received wide media coverage that included the press and a large press conference, TIMSS 2011 results were only commented on in several newspapers.

International large-scale assessment results have been used by policy makers and education leaders to improve education quality in Jordan. For example, they have been used to track the impact of education reforms on student achievement levels (including informing interventions such as ERfKE), as well as inform curriculum improvement, teacher training programs, and teacher guidelines.

However, there is no evidence that the use of international large-scale assessment data has positively impacted student achievement levels. The use of international assessment information is neither systematic nor structured, with little analysis conducted on the results. Thus, while interventions that result from international large-scale assessment data take place, they are not implemented on a large scale and are insufficient to impact overall student achievement in the country. Recently, international large-scale assessment results revealed a marked drop in student performance; policy makers have since taken concrete actions to study the reasons behind the decline.
Following the administration of each such assessment, NCHRD prepares a national report that contains a full analysis of results, as well as policy briefs targeted to policy makers. The final national report is made publically available at the National Library. International large-scale assessment results are also shared during a half-day meeting with stakeholders, including representatives of Ministry of Education Directorates (planning, examinations and tests, curriculum, education training center, DCU), university professors, the Ministry of Labor, and other ministerial staff who are interested in attending.

Together with the relevant directorates of the Ministry of Education, NCHRD also prepares hard copies of the international report and publishes the raw data in CD form. The national and international reports, as well as the policy brief, are shared with other ministries, universities, and stakeholder groups.

Uses of assessment results
High-level officials, including Her Majesty Queen Rania, attach high importance to international large-scale assessments. The results of these assessments significantly influence policy making at the national level, signaling the extent to which reforms have been successful. Average student scores, as well as the ranking of Jordan against other countries, improved in each international assessment through 2009. For example, on TIMSS 1999, Jordan ranked 30th in science and 33rd in math out of 38 participating countries; on TIMSS 2007, Jordan ranked 21st in science and 31st in math out of 50 participating countries. In addition, the average student score increased from 450 in 1999 to 483 in 2007. The results of such assessments have been used to monitor trends at the country level and compare the performance of Jordanian students with students of other countries.

Traditionally, as a follow-up to TIMSS findings, NCHRD works closely with the Ministry of Education to prepare manuals for math and science teachers to help them improve their practices and methodologies. The manuals are intended to assist them in addressing common errors made by students. These manuals are used in teacher training programs. Under ERfKE II, the Ministry of Education also began using PISA and TIMSS results to refine the curricula. Studies have also been undertaken that compare Jordanian education policies to those of high-achieving countries, such as Singapore, Taiwan, and Japan.

Consistent with NAfKE, the results of TIMSS 2011 and PISA 2012 showed a serious decline in student performance compared to previous years, both in terms of Jordan’s ranking and the average student score, raising concerns among education policy makers and their partners in Jordan. In response, a high-level committee was established within the Ministry of Education. In addition, a full analysis of the obstacles affecting student performance has been conducted that considered teaching, learning, and socioeconomic factors.

Accordingly, an action plan was developed to address the main issues that contributed to the decline in performance. The action plan is divided into six
components: (i) analysis of results/decline factors; (ii) awareness raising; (iii) time allocated to teaching; (iv) in-service teacher training; (v) curriculum; and (vi) assessments (MoE, 2013a). For each component, a detailed list of activities and steps to be taken by the relevant entity or directorate are outlined. For instance, as part of awareness raising among teachers, schools, and students, directorates are urged to launch a website on international assessments. Implementation of this action plan is currently underway.

6. Drivers for Change

Multiple drivers contributed to the development of the student assessment system in Jordan, including the economic goal of building a knowledge economy, the national agenda, ERfKE, and donor agencies.

Education reforms in the country have been greatly accelerated under His Majesty King Abdullah II, guided by his vision of creating a skills-, knowledge- and innovation-driven economy. In early 2000, he called for the “remodeling” of the education system as a critical step toward Jordan’s vision of becoming a regional information technology hub and fully entering into the global economy.7 The National Agenda (Government of Jordan 2005) is the first major attempt to translate His Majesty’s vision into a coherent agenda for Jordan’s socioeconomic and political development, prompting the beginning of an era of change in the country. The agenda identified eight pillars as national priorities,8 among which education is a major component of two:

- education, higher education, scientific research, and innovation; and
- employment support and vocational training.

Jordan has undertaken several structural reforms to align the education system with the overall direction of the National Agenda, with ERfKE representing the most comprehensive reform.

As noted earlier, ERfKE is a 10-year, two-part, multidonor education sector reform program designed to deliver on the country’s 2002 vision for its education system (World Bank 2009) by re-engineering it to produce relevant skills for the knowledge economy.9 Recognizing that the benefits of education accrue to

---

8 The eight priorities are: political development and inclusion; justice and legislation; investment development; financial services and fiscal reform; employment support and vocational training; social welfare; education, higher education, scientific research, and innovation; and infrastructure upgrades.
9 The 2002 Vision Forum for the Future of Education in Jordan states: “The Hashemite Kingdom of Jordan has the quality competitive human resource development systems that provide all people with lifelong learning experiences relevant to their current and future
society only when learning occurs (OECD 2010), ERfKE focused on enhancing the way in which learning is measured as part of a curricular reform and reforms of the teaching and learning process.

ERfKE provided leadership, direction, and financing so that student assessment activities could shift from being implemented on an ad hoc to a systematic basis. The program specifically focused on strengthening the monitoring and evaluation of educational outcomes at all levels of the education system.

In addition, the establishment of NCHRD was a crucial part of developing the student assessment system, allowing a specialized body external to the Ministry of Education to lead external evaluations and large-scale assessments. Along with MoE, NCHRD developed a framework for student assessments and created an accountability mechanism in the Ministry of Education for monitoring education results. The introduction of the joint monitoring and evaluation framework has increased transparency and created an accountability system at the school level.

Another key driver of the improved student assessment system was the dissemination strategy for TIMSS, PISA, and NAfKE results. The improved performance of students on international large-scale assessments from 1999 to 2007 resulted in less attention being paid to the use of these results because a positive trend was seen as proof that there were no problematic issues in the education system. The decline in student performance compared to previous years on NAfKE 2011, TIMSS 2011, and PISA 2012 pushed the Ministry of Education to challenge existing methods of using and disseminating the results of these assessments.

MoE and NCHRD collaborated to conduct an in-depth analysis of international large-scale assessment results, seeking to understand the reasons for the decline in student scores. MoE then formed a committee to take concrete actions to improve performance. This committee meets on a regular basis to prepare recommendations to improve student scores on international assessments. For example, the committee has recommended making teachers, students, and parents aware of the importance of student assessments. Although they had previously been regarded as primarily the concern of central agencies and policy makers, Jordan’s involvement in international assessments and the use and dissemination of assessment data has shifted the perception of policy makers, who now recognize the importance of engaging the entire school community in addressing student performance.

International influences also helped shape the use of student assessment data. The education system has, for example, received substantial funding from international donors through ERfKE. Partners in educational reforms, such as the

needs in order to respond to and stimulate sustained economic development through an educated population and an educated workforce” (as cited in World Bank 2009).
World Bank, U.S. Agency for International Development, and Canadian International Development Agency, have stressed the importance of creating a sustained assessment framework and have pushed for greater transparency of learning outcomes, both of which make responsible entities more accountable for results. The MoE was responsible for reporting regularly to donors on the performance of the project, and on the education system overall, using the indicators established by the ERfKE M&E framework.

7. Discussion

Jordan has gradually built a hybrid assessment system that encompasses the four main types of assessments used by the majority of education systems in the world: classroom assessments, examinations, and national and international large-scale assessments. Beginning with the Tawjihi and traditional classroom assessment activities, the system has evolved to include updated and forward-thinking classroom assessment practices, as well as national and international large-scale assessments that influence the direction of educational reforms in the country. Figure 7 summarizes the key features of student assessment programs in Jordan.

While the assessment system in Jordan has developed to include various components, the linkages across the different assessments remain weak. Other than applying design aspects of international assessments in developing national assessment tools, various assessment activities are conducted and managed in isolation from one another, and coordination is limited among responsible agencies. This is partly due to the lack of one entity assigned to oversee student assessments as a system. In order to address this gap, NCHRD, with the technical assistance of the Monitoring and Evaluation Project (MEP),10 is in the process of mapping student assessment tools. The purpose of this mapping exercise is to look at and compare the different assessments and the tools used, discern data trends, and compare both the way assessments are used and the way they impact pedagogical practices.

10 USAID funded-project.
### Figure 7. Summary of Assessment Programs in Jordan

<table>
<thead>
<tr>
<th>Purpose</th>
<th>NAfKE (since 2006)</th>
<th>National Test (since 2000, reformed in 2004)</th>
<th>&quot;Tawjih&quot; (since 1962)</th>
<th>Classroom assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Informing policy interventions and measuring impact</td>
<td>Providing pedagogical support</td>
<td>Certifying students and selecting them into post-secondary education</td>
<td>Informing pedagogy at the classroom level, supporting student learning, diagnosing issues, tracking students in secondary education</td>
</tr>
<tr>
<td>Audience</td>
<td>Decision makers and educational partners</td>
<td>School principal, pedagogical coordinators, and teachers</td>
<td>Students and parents</td>
<td>Teachers, school principals, pedagogical coordinators, students, and parents</td>
</tr>
<tr>
<td>Population</td>
<td>Sample of students in grades 5, 9, and 11</td>
<td>Students in grades 4, 8, and 10; however, each year students from only one of the three grades are assessed</td>
<td>12th-grade students</td>
<td>Students in the classroom</td>
</tr>
<tr>
<td>What is assessed</td>
<td>Math, science, and literacy (knowledge-economy skills in particular)</td>
<td>Arabic, mathematics, science, culture, English</td>
<td>Five common courses (Arabic, Islamic culture, English, general culture, computers) and specialized courses by individual stream</td>
<td>Knowledge of curriculum and skills (both academic and knowledge-economy skills)</td>
</tr>
<tr>
<td>Frequency</td>
<td>2 to 4 times every five years</td>
<td>One grade assessed annually</td>
<td>Annually</td>
<td>Routinely/daily</td>
</tr>
<tr>
<td>Responsible entity</td>
<td>NCHRD</td>
<td>MoE (Directorate of Examinations and Tests)</td>
<td>MoE (Directorate of Examinations and Tests)</td>
<td>MoE</td>
</tr>
</tbody>
</table>

Source: Authors.

There is still room to improve the use of student assessment data in Jordan. It remains to be determined, for example, to what extent stakeholders read the reports and policy briefs that are disseminated. A study undertaken by World Education, Inc. and NCHRD on how data from different assessments (including PISA, TIMSS and NAfKE) are used showed that most targeted MoE staff did not read any of the reports and policy briefs produced under ERfKE II (WEI & NCHRD 2013) (figure 8).
As the figure makes clear, executive summaries are more likely to be read than the full reports. Reading the findings of assessment studies is a prerequisite for the potential use of the data. In addition, results did not differ among staff working in the central offices of MoE and those working in the governorates (figure 9).
In addition to examining the extent to which stakeholders read assessment findings in different formats, the same study sought to learn how policy makers used such reports. The findings revealed that among staff interviewed, NAfKE and PISA data were the least utilized and Classroom Observation Study data were highly utilized (figure 10).\textsuperscript{11} One of the reasons for this underutilization could be that staff are unable to understand the content of the reports, and that the quality and relevance of the statistics, data, concluding remarks, and policy implications presented in the reports needs to be improved. At the same time, the study found that there is a positive correlation between participation in the development of assessment or evaluation tools and their Terms of Reference, on one hand, and stakeholders’ opinion of the usefulness of evaluations, on the other.

**Figure 10. Were Findings Presented in Reports Useful in the Education Decision-Making Process?**

![Figure 10. Were Findings Presented in Reports Useful in the Education Decision-Making Process?](image)


Student assessment activities can be a major driver for the introduction of both accountability and incentives. Many countries use assessment programs to hold schools accountable, which can trigger changes at the school level; however, Jordan has not yet used its assessment program for this purpose.

\textsuperscript{11} The Classroom Observation Study is one of the external evaluations conducted by NCHRD that uses classroom observation methods to measure the extent to which the ERfKE program contributes to enhanced teacher performance in the classroom.
8. Lessons Learned

Jordan has gone through several stages in building its current student assessment system. Countries aiming to build a hybrid system and strengthen the use and dissemination of assessment data can learn a number of lessons from Jordan’s experience.

Set clear objectives. It is very important to identify the purposes of an assessment, based on a clear vision of what will be assessed and why; ultimately, this vision will influence the use of assessment results. Ad-hoc activities without a clear direction or framework will not lead to desired changes or improvements in the education system. Political influence and leadership is critical in setting objectives and overall direction. Jordan’s National Agenda and broader educational reforms have provided the overall direction for its student assessment system by pushing for better monitoring of results and effective accountability mechanisms.

Gradually implement student assessments. While it is important to look at best practices and discern lessons from other countries, changes to a national student assessment system should be introduced gradually. This is in part due to the capacity of the implementing body and the country to absorb assessment information. Jordan did not introduce complex strategies for student assessment suddenly; the system was built gradually. At every stage, the monitoring and evaluation framework was reviewed and small adjustments were made to both the design of assessment instruments and the dissemination and use of assessment data. For example, a team from NCHRD is responsible for reviewing the instruments and the implementation process following each round of national assessments.

Even the reporting of assessment results has evolved over time. When Jordan participated in an international assessment for the first time, the national report consisted only of results. Today, the national report presents a detailed analysis of results and a discussion of the socioeconomic factors affecting those results. These improvements were the result of a change in priorities at the country level that pushed for better accountability, monitoring, and capacity building for the NCHRD implementation team.

Create an external organization. An external organization such as the NCHRD ensures impartial assessment of the strengths and weaknesses of a country’s education system, independent of political pressure and validated by the results of other internal monitoring and evaluation activities. Jordan benefited from the establishment of NCHRD because the Ministry of Education had limited capacity to lead large-scale assessments, and because NCHRD brings an independent perspective to educational research and human development activities at the
national level. In addition, having another body that, along with Ministry of Education, assesses the progress of the education system as a whole in a transparent manner has helped improve the accountability of the ministry’s departments, particularly those responsible for monitoring and evaluation activities. Finally, the exposure of NCHRD to international best practices and its sole focus on research and assessment enabled it to incorporate aspects of international assessments into national assessment instruments so as to further improve the latter.

Establish a clear overall framework for assessment activities. The ERfKE monitoring and evaluation framework is the only document that clearly identifies the main assessment parameters for the Jordanian education system; it also provides a clear roadmap for assessment activities and sets broader educational goals. The division of labor between NCHRD and MoE—and within the MoE (e.g., the Directorates of Examinations and Tests and M&E)—was critical to the effective implementation of ERfKE M&E activities, including student assessments. Prior to the establishment of this framework, the roles and responsibilities of different bodies in managing assessment activities and disseminating results were not clear, which led to redundancy and inefficiency.

Conduct reviews and evaluations. As important as it is to have a clear framework, it is equally important to continuously review whether assessment tools are being utilized effectively for their intended purposes. As noted earlier, the first study that looked at this issue (WEI and NCHRD 2013) showed that many target audiences do not read assessment data, and that the extent to which the Ministry of Education incorporated these data into its policies and plans was very limited. The study also examined whether studies that measured ERfKE results were meeting their objectives. These types of reviews are very important to ensure continuous improvement of the framework, allowing for corrective action at the right time.

Have a clear dissemination strategy. A clear strategy for disseminating assessment results helps orient assessment activities. The strategy must identify what will be disseminated, to whom, and who will be responsible for dissemination, ensuring that results are communicated and used in a timely manner. A common problem in Jordan was that dissemination was ad-hoc and not targeted; in addition, long reports often included many details and the information they contained was sometimes not necessarily utilized or interpreted correctly.

The aforementioned framework therefore identified which entities were responsible for dissemination of which assessment data. As a result, NCHRD started preparing different materials that targeted different audiences, such as a policy brief specifically developed for policy makers, who usually do not read long reports and may only need a summary. For educators and practitioners, the
center leads in-depth presentations during a workshop hosted by MoE. It is also important to identify who should disseminate assessment results. While NCHRD administers international assessments such as TIMSS and PISA and analyzes their results, MoE is responsible for communicating these results to its staff. These different dissemination strategies were very useful for expanding the distribution of assessment information and strengthening the culture of assessment.

**Train teachers in classroom assessment strategies.** It is important to ensure that teachers are not overloaded with too many classroom assessment strategies, but are instead trained on the most effective ones. For example, when the assessment strategy of 2004 first introduced such new concepts as “authentic assessment,” teachers were not sufficiently trained on how to use the methodologies associated with such assessment activities in the classroom. After several years it became evident that teachers were not applying the new strategies in their teaching, but rather, were resorting to conventional assessment activities. The Ministry of Education thus recognized the need to conduct extensive training for both teacher supervisors and teachers to ensure that the assessment reforms introduced at the national level were applied at the school level and implemented consistently across schools.

**Create linkages.** While Jordan’s student assessment system developed substantially over the past two decades as various components of the system were introduced, the linkages across assessment activities were not developed equally. There seems to be a disconnect between the different assessments, as well as among their implementing entities. Creating linkages could improve various aspects of assessment tools, their implementation, and use. Sharing lessons and streamlining activities such as dissemination could increase both efficiency and effectiveness. So far, the only area where linkages have been created is in the design of instruments. Specifically, having NCHRD as the entity that administers both international assessments and NAfKE allowed for these linkages to be created. Thus, NAfKE instruments are designed and updated in alignment with international standards. In addition, the National Test was also recently adapted to meet some of the standards that were applied to NAfKE and aligned with the latter assessment.

**Establish a clear baseline.** Having a baseline is a prerequisite for measurement and monitoring of progress and trends. Both NAfKE and the National Test failed to establish a clear baseline against which to measure. The changes and irregularities in the design of the National Test and the way in which its results were presented made it very difficult to discern trends over the years. As for NAfKE, the baseline was also not clearly set or identified, particularly because the assessment is linked to ERfKE and its phases.
Conclusion

Against a background of broader education reforms, Jordan has taken important steps to: (i) improve the student assessment system so that it plays a role in both measuring learning outcomes and supporting the learning process; (ii) enhance the effectiveness of its assessment tools; and (iii) strengthen the use of assessment data for its intended purpose. Positive results have been achieved so far, and Jordan continues to improve how it uses assessment data, especially following the recent decline in student performance on TIMSS, PISA, and NAfKE.

The experience of Jordan shows that in the process of building a student assessment system, leadership is key. The broader vision of Jordan and the ERfKE program played a critical role in improving student assessment activities and their use. ERfKE has provided leadership, direction, and financing for student assessment activities. It has helped the country shift from ad-hoc to systematic implementation of learning assessments by strengthening monitoring and evaluation of learning outcomes at all levels of education.
References


List of papers published in the current Working Paper series


Student assessment has been increasingly recognized as playing a key role in monitoring and improving student learning and achievement levels. This case study discusses the important steps that Jordan has taken to strengthen its student assessment system and ensure that the information gathered from classroom assessments, examinations, and large-scale assessment activities supports student learning. Countries seeking to strengthen the utilization and dissemination of student assessment data can extract the following lessons from Jordan’s experience: they should (i) set clear objectives for their assessment system, based on a well-defined vision of what is being assessed and why; (ii) entrust an external entity with managing system-level assessments in order to ensure impartial evaluation of the strengths and weaknesses of the education system; and (iii) introduce change gradually, coupled with adequate training on new assessment strategies and methods.

Osamha Obeidat, Inspection Program Manager at Abu Dhabi Education Council

Zaina Dawani, education consultant in the World Bank’s Middle East and North Africa region